

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2006, 12:43:40 ; Search time 200 Seconds

(without alignments)
1682.556 Million cell updates/sec

Title: US-10-696-261-13

Perfect score: 3989 1 MADGYLPPWLEDNISEGIR.....NNGLYTEPRPIGRYLRPL 736

Sequence: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_8:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003s:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

9: geneseqp2005s:*

10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3989	100.0	736	3	AAV71167
2	3989	100.0	736	6	ABBB80232
3	3989	100.0	736	7	ABR62762
4	3989	100.0	736	7	ADZ76565
5	3989	100.0	736	9	ADV70294
6	3989	100.0	736	9	ADZ27069
7	3989	100.0	736	9	ADZ27012
8	3985	99.9	735	9	ADV67506
9	3985	99.9	736	9	ADZ27007
10	3963	99.3	736	4	AAAB59847
11	3963	99.3	736	7	ADZ76566
12	3963	99.3	736	9	ADV70293
13	3963	99.3	736	9	ADZ27070
14	3897.5	97.7	737	9	ADZ27086
15	3883.5	97.4	737	9	ADZ27010
16	3883	97.3	736	9	ADZ27074
17	3873	97.1	736	9	ADZ27008
18	3853	96.6	736	9	ADZ27009
19	3511	88.0	736	4	AAAB59846
20	3494	87.6	736	4	AAAB59845
21	3494	87.6	736	6	ABBB80233
22	3494	87.6	736	6	ABR62763
23	3494	87.6	736	7	ADZ76572

24	3494	87.6	736	9	ADV67507
25	3494	87.6	736	9	ADZ27067
26	3483.5	87.3	737	9	ADZ26959
27	3476	87.1	736	9	ADZ27085
28	3473	87.1	736	7	ADZ76598
29	3469	87.0	738	9	ADZ26960
30	3467	86.9	736	6	ABBB80229
31	3467	86.9	736	7	ABR62764
32	3467	86.9	736	7	ADZ76601
33	3467	86.9	736	9	ADV67508
34	3464.5	86.9	737	9	ADZ27080
35	3464	86.8	736	7	ADZ76597
36	3460	86.7	736	7	ADZ76599
37	3458.5	86.7	737	9	ADZ26963
38	3457	86.7	736	7	ADZ76600
39	3456.5	86.7	735	9	ADZ27034
40	3456.5	86.7	737	9	ADZ26966
41	3456.5	86.7	737	9	ADZ26962
42	3456	86.6	738	9	ADZ26951
43	3455.5	86.6	737	9	ADZ26965
44	3453.5	86.6	735	9	ADZ26996
45	3453.5	86.6	735	9	ADZ26992

ALIGNMENTS

RESULT 1
ID AAV71167 standard; protein; 736 AA.
AAV71167
AC AAV71167;
DT 08-SEP-2000 (first entry)
XX
DE Adeno-associated virus serotype 1 capsid protein VP1.
XX
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
KM cap protein; recombinant viral vector; gene delivery; gene therapy;
KN vaccine; transgene; VP1.
XX
OS Adeno-associated virus 1.
XX
PN WO200028061-A2.
XX
PD 18-MAY-2000.
XX
PF 02-NOV-1999; 99MO-US025694.
XX
PR 05-NOV-1998; 98US-0107114P.
XX
PA (UNIPROT) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Xiao W;
XX
XX WPI; 2000-376571/32.
DR N-PSDB; AAD00772, AAD00777.
XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for
PT preparation of medicament for delivery of a transgene to a host.
XX
PS Claim 7; Page 87-90; 108pp; English.
XX
CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3. The
CC AAV-1 sequence or its fragments particularly ITRs, rep and cap coding
CC regions, are useful in production of recombinant viral vectors for gene
CC delivery. These vectors can be used as gene therapy vectors, vaccine
CC vectors or antisense delivery vectors. The AAV-1 does not induce the
CC formation of neutralising antibodies specific to any serotype of AAV

CC hence is useful for transfecting host cells, and in the preparation of a
 CC medicament for the delivery of transgene to a host. The present sequence
 CC is an AAV-1 cap protein VP1 which is useful in the production of
 CC recombinant viral vector for gene delivery

XX
 SO Sequence 736 AA:

Query Match 100.0%; Score 3989; DB 3; Length 736;
 Best Local Similarity 100.0%; Pred. No. 5.2e-311;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGTLPLWLENDLSEGRREWMWDLKPGAPKPKANQOKODGRLVPGYKYLGFPGGLD 60
 DB 1 MAADGTLPLWLENDLSEGRREWMWDLKPGAPKPKANQOKODGRLVPGYKYLGFPGGLD 60
 QY 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAEFOERLOEDTSFGNUGRAVFQ 120
 DB 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAEFOERLOEDTSFGNUGRAVFQ 120
 QY 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGOQPAKKRLNFGQTGDS 180
 DB 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGOQPAKKRLNFGQTGDS 180
 QY 181 SVDDPQPLGEPPTPAVGPPTMASGGGAPMADNNEGADVGNASGNWCHDSTWLGDRVI 240
 DB 181 SVDDPQPLGEPPTPAVGPPTMASGGGAPMADNNEGADVGNASGNWCHDSTWLGDRVI 240
 QY 241 TTSRTWALPTVNNHLYKQISSASTGASNDNHFGYSTPMGYDFNRFHGFSPRDMQRL 300
 DB 241 TTSRTWALPTVNNHLYKQISSASTGASNDNHFGYSTPMGYDFNRFHGFSPRDMQRL 300
 QY 301 INNNWGRPRKRLNPKLFNIOVKVTTNDGVTIANNLTSVQVPSDEYOLPYVLGSAHQ 360
 DB 301 INNNWGRPRKRLNPKLFNIOVKVTTNDGVTIANNLTSVQVPSDEYOLPYVLGSAHQ 360
 QY 361 GCLPPFPADVFMIPQGYLTLNNGSQAVGRSFFCYCLEYFPSOMLRTGNNTFSYTFEEVP 420
 DB 361 GCLPPFPADVFMIPQGYLTLNNGSQAVGRSFFCYCLEYFPSOMLRTGNNTFSYTFEEVP 420
 QY 421 FHSSYAHOSGLDRLNMPLDIOYLYLNRTONOGSAQNKDLPFRSGPAGMSVQPKWMLP 480
 DB 421 FHSSYAHOSGLDRLNMPLDIOYLYLNRTONOGSAQNKDLPFRSGPAGMSVQPKWMLP 480
 QY 481 GPCYRQORVSKITDNNNSNFTWTGASKYNLNGRESIINPGTMAASHKDEDEKFFPMGSV 540
 DB 481 GPCYRQORVSKITDNNNSNFTWTGASKYNLNGRESIINPGTMAASHKDEDEKFFPMGSV 540
 QY 541 MIFGKESAGASNTALDNVMTDEBEIKATNPVATERGTVAVNPOSSSTDPAIGDVFAMG 600
 DB 541 MIFGKESAGASNTALDNVMTDEBEIKATNPVATERGTVAVNPOSSSTDPAIGDVFAMG 600
 QY 601 ALPGMWQDRDYYLQGFIMAKIPHTDGHFHPSPLMGFGKPNPPQILINNTVPANPPA 660
 DB 601 ALPGMWQDRDYYLQGFIMAKIPHTDGHFHPSPLMGFGKPNPPQILINNTVPANPPA 660
 QY 661 EBSATKPASTTQYSTQCVSEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 720
 DB 661 EBSATKPASTTQYSTQCVSEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 720
 QY 721 YTEPRPIGRYTLRPL 736
 DB 721 YTEPRPIGRYTLRPL 736

RESULT 2
 ABB80232
 ID ABB80232 standard; protein; 736 AA.

XX ABB80232;
 XX 20-NOV-2003 (first entry)
 DT
 XX
 DE AAV1 vp1 protein.

XX
 KW Adeno-associated virus; AAV; serotype 9; rep; cap; vp1; vp2; vp3;
 KW splice variant; transgene.

XX
 OS Adeno associated virus serotype 1.

XX
 PN W02003052052-A2.

XX
 PD 26-JUN-2003.

XX
 PF 12-NOV-2002; 2002W0-US033631.

XX
 PR 17-DEC-2001; 2001US-0341150P.

XX
 PR 05-JUN-2002; 2002US-0386132P.

XX
 PA (TYPE-) UNIV PENNSYLVANIA.

XX
 PI Gao G, Wilson JM, Alvira M;

XX
 DR WPI; 2003-523523/49.

XX
 PT New isolated adeno-associated virus (AAV) comprising an AAV9 capsid,
 useful for preparing a medicament for delivering a transgene to a cell.

XX
 PS Disclosure; Fig 2; 76pp; English.

XX
 CC The sequences given in ABB80231-34 represent vp1 proteins derived from
 CC various adeno-associated virus (AAV) serotypes. These sequences were used
 CC in the scope of the invention for comparison with the cap protein derived
 CC from AAV serotype 9. The AAV capsid comprises three proteins VP1, VP2 and
 CC VP3, which are alternative splice variants. The AAV or the nucleic acid
 CC molecule is useful for preparing a medicament for delivering a transgene
 CC to a cell

XX
 SO Sequence 736 AA:

Query Match 100.0%; Score 3989; DB 6; Length 736;
 Best Local Similarity 100.0%; Pred. No. 5.2e-311;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGTLPLWLENDLSEGRREWMWDLKPGAPKPKANQOKODGRLVPGYKYLGFPGGLD 60
 DB 1 MAADGTLPLWLENDLSEGRREWMWDLKPGAPKPKANQOKODGRLVPGYKYLGFPGGLD 60
 QY 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAEFOERLOEDTSFGNUGRAVFQ 120
 DB 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAEFOERLOEDTSFGNUGRAVFQ 120
 QY 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGOQPAKKRLNFGQTGDS 180
 DB 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGOQPAKKRLNFGQTGDS 180
 QY 181 SVDDPQPLGEPPTPAVGPPTMASGGGAPMADNNEGADVGNASGNWCHDSTWLGDRVI 240
 DB 181 SVDDPQPLGEPPTPAVGPPTMASGGGAPMADNNEGADVGNASGNWCHDSTWLGDRVI 240
 QY 241 TTSRTWALPTVNNHLYKQISSASTGASNDNHFGYSTPMGYDFNRFHGFSPRDMQRL 300
 DB 241 TTSRTWALPTVNNHLYKQISSASTGASNDNHFGYSTPMGYDFNRFHGFSPRDMQRL 300
 QY 301 INNNWGRPRKRLNPKLFNIOVKVTTNDGVTIANNLTSVQVPSDEYOLPYVLGSAHQ 360
 DB 301 INNNWGRPRKRLNPKLFNIOVKVTTNDGVTIANNLTSVQVPSDEYOLPYVLGSAHQ 360
 QY 361 GCLPPFPADVFMIPQGYLTLNNGSQAVGRSFFCYCLEYFPSOMLRTGNNTFSYTFEEVP 420
 DB 361 GCLPPFPADVFMIPQGYLTLNNGSQAVGRSFFCYCLEYFPSOMLRTGNNTFSYTFEEVP 420
 QY 421 FHSSYAHOSGLDRLNMPLDIOYLYLNRTONOGSAQNKDLPFRSGPAGMSVQPKWMLP 480
 DB 421 FHSSYAHOSGLDRLNMPLDIOYLYLNRTONOGSAQNKDLPFRSGPAGMSVQPKWMLP 480
 QY 481 GPCYRQORVSKITDNNNSNFTWTGASKYNLNGRESIINPGTMAASHKDEDEKFFPMGSV 540

```

Db      481 GPCRRQGRVSKTKTDNNNSNFTWTGASKYLNNGRESIIINPTAASHKDEDEKFFPMGSV 540
Qy      541 MIFGESAGASNTALDVMITDEBEIKATNPVATERFGTVAANFQSSSTDPATGDVAMG 600
Db      541 MIFGESAGASNTALDVMITDEBEIKATNPVATERFGTVAANFQSSSTDPATGDVAMG 600
Qy      601 ALPGMWQDRDVLVYLGQIPMAKIPHTDGHFHPSPLMGFGKKNPPQILIKNTVPANPPA 660
Db      601 ALPGMWQDRDVLVYLGQIPMAKIPHTDGHFHPSPLMGFGKKNPPQILIKNTVPANPPA 660
Qy      661 EFSATKASFTITQYSTGQVSEIEMELQKNSKRNPEVOYTSNYAKSANVDFTVDNGL 720
Db      661 EFSATKASFTITQYSTGQVSEIEMELQKNSKRNPEVOYTSNYAKSANVDFTVDNGL 720
Qy      721 YTEBRPIGTRYLTRPL 736
Db      721 YTEBRPIGTRYLTRPL 736

```

RESULT 3
ABR62762
ID ABR62762 standard; protein; 736 AA.

```

XX      AC ABR62762;
XX      DT 06-NOV-2003 (first entry)
XX      DE Adeno associated virus 1 capsid protein vp1.
XX      KM AAV; AAV1; capsid; vector; gene therapy; antisense therapy; vaccine.
XX      OS Adeno associated virus.
XX      PN MO2003052051-A2.
XX      PD 26-JUN-2003.
XX      PF 12-NOV-2002; 2002MO-US033630.
XX      PR 17-DEC-2001; 2001US-0341151P.
XX      PR 01-MAY-2002; 2002US-0377133P.
XX      PR 05-JUN-2002; 2002US-0386122P.
XX      PA (UYPE-) UNIV PENNSYLVANIA.
XX      PI Gao G, Wilson JM, Alvirra M;
XX      PS WPI; 2003-523522/49.
XX      PT New adeno-associated virus (AAV) comprising an AAV capsid having an amino
XX      PT acid sequence of AAV8, useful for preparing a medicament for delivery of
XX      PT a transgene to a cell and for treating cystic fibrosis or hemophilia B.
XX      PS Disclosure; Fig 2A-C; 82bp; English.
XX
XX      CC The present sequence is the protein sequence of the vp1 capsid protein of
XX      CC adeno associated virus serotype 1 (AAV1). The invention provides the
XX      CC nucleic acid and amino acid sequences of novel AAV8 and fragments of
XX      CC these sequences. Each of these fragments may be used in a variety of
XX      CC vector systems and host cells. Among the desirable fragments are the cap
XX      CC proteins, including vp1, vp2, vp3 and hypervariable regions, the rep
XX      CC proteins, including rep78, rep68, rep52 and rep40, and the sequences
XX      CC encoding these proteins. The fragments may be used alone, in combination
XX      CC with other AAV8 sequences or in combination with elements from other AAV
XX      CC or non-AAV8 sequences in the production of recombinant AAV and for
XX      CC use as antisense delivery vectors, gene therapy vectors or vaccine
XX      CC vectors. A claimed molecule comprises a cap protein of a functional AAV
XX      CC rep gene from a serotype selected from AAV1, AAV2, AAV3, AAV4, AAV5 and
XX      CC AAV6
XX      SO Sequence 736 AA.

```

Query Match 100.0%; Score 3989; DB 7; Length 736;
Best Local Similarity 100.0%; Pred. No. 5.2e-311;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MAADGYLPDLIEDNLSEGIEMWDLKPGAKPKANQOKODGRGLVPGYKYLGFENGLD 60
Db      1 MAADGYLPDLIEDNLSEGIEMWDLKPGAKPKANQOKODGRGLVPGYKYLGFENGLD 60
Qy      61 KGEPPNADAAALHEHKAAYDQQLKAGDPYLRYNHADAERQERLOEPTSGGNIGRAVFO 120
Db      61 KGEPPNADAAALHEHKAAYDQQLKAGDPYLRYNHADAERQERLOEPTSGGNIGRAVFO 120
Qy      121 AKKRVLEPLGLVEGATAPGKKRPVQSPQOEPPSSGIGTGQOPAKKLNFGQTDSE 180
Db      121 AKKRVLEPLGLVEGATAPGKKRPVQSPQOEPPSSGIGTGQOPAKKLNFGQTDSE 180
Qy      181 SVDPQPLGEPPTATPAVAGPTTMAAGGCAEMADNNEGADGVGNAGNMHCDSTWLAGRVI 240
Db      181 SVDPQPLGEPPTATPAVAGPTTMAAGGCAEMADNNEGADGVGNAGNMHCDSTWLAGRVI 240
Qy      241 TTSRTWALPTYNHLYKQISSASTGASDNDHYFGYSTPGYDFENRHHFSPRDMQRL 300
Db      241 TTSRTWALPTYNHLYKQISSASTGASDNDHYFGYSTPGYDFENRHHFSPRDMQRL 300
Qy      301 INNMGFRPRKLNPKLFNIOKVEVTNDGYTTANNLTSTVOVPSDSEYOLPYLGSAAHQ 360
Db      301 INNMGFRPRKLNPKLFNIOKVEVTNDGYTTANNLTSTVOVPSDSEYOLPYLGSAAHQ 360
Qy      361 GCLPPFPADVFMIPOYGYLTINNGSQAVGRSSFYCLEYFPSONMLRTGNFTFSYTFEEVP 420
Db      361 GCLPPFPADVFMIPOYGYLTINNGSQAVGRSSFYCLEYFPSONMLRTGNFTFSYTFEEVP 420
Qy      421 FHSYSAHSQGLDRIMNPLIDQYLYLINRTQNGSQAQNKDLFSGSGSPAGMSVQPKWLP 480
Db      421 FHSYSAHSQGLDRIMNPLIDQYLYLINRTQNGSQAQNKDLFSGSGSPAGMSVQPKWLP 480
Qy      481 GPCYRQGRVSKTKTDNNNSNFTWTGASKYLNNGRESIIINPTAASHKDEDEKFFPMGSV 540
Db      481 GPCYRQGRVSKTKTDNNNSNFTWTGASKYLNNGRESIIINPTAASHKDEDEKFFPMGSV 540
Qy      541 MIFGESAGASNTALDVMITDEBEIKATNPVATERFGTVAANFQSSSTDPATGDVAMG 600
Db      541 MIFGESAGASNTALDVMITDEBEIKATNPVATERFGTVAANFQSSSTDPATGDVAMG 600
Qy      601 ALPGMWQDRDVLVYLGQIPMAKIPHTDGHFHPSPLMGFGKKNPPQILIKNTVPANPPA 660
Db      601 ALPGMWQDRDVLVYLGQIPMAKIPHTDGHFHPSPLMGFGKKNPPQILIKNTVPANPPA 660
Qy      661 EFSATKASFTITQYSTGQVSEIEMELQKNSKRNPEVOYTSNYAKSANVDFTVDNGL 720
Db      661 EFSATKASFTITQYSTGQVSEIEMELQKNSKRNPEVOYTSNYAKSANVDFTVDNGL 720
Qy      721 YTEBRPIGTRYLTRPL 736
Db      721 YTEBRPIGTRYLTRPL 736

```

RESULT 4
ADE76565
ID ADE76565 standard; protein; 736 AA.
XX AC ADE76565;
XX DT 29-JAN-2004 (first entry)
XX DE Adeno-associated virus (AAV) related protein, SEQ ID No 64.
XX KM adeno-associated virus; AAV; cytotatic; antiproliferative; antineoplastic;
XX KM antiarthritic; neuroprotective; antidiabetic; antithyroid;
XX KM dermatological; antiinflammatory; gene therapy; vaccine;
XX KM hyperproliferative; cancer; psoriasis; autoimmune disease;
XX KM rheumatoid arthritis; multiple sclerosis; diabetes;
XX KM autoimmune thyroiditis; scleroderma; Crohn's disease.

XX Unidentified.
 OS
 XX
 PN EP1310571-A2.
 XX
 PD 14-MAY-2003.
 XX
 PF 12-NOV-2002; 2002EP-00257826.
 XX
 PR 13-NOV-2001; 2001US-0350607P.
 PR 17-DEC-2001; 2001US-0341117P.
 PR 01-MAY-2002; 2002US-0377066P.
 PR 05-JUN-2002; 2002US-0386675P.
 XX
 PA (UYPE-) UNIV PENNSYLVANIA.
 XX
 PI Gao G, Wilson JM, Alvira M;
 XX
 DR WPI; 2003-450984/43.
 XX
 PT Detecting adeno-associated virus sequences in a sample, useful for e.g.
 PT preventing or treating hyperproliferative or autoimmune diseases,
 PT comprises subjecting a sample having a DNA to amplification via
 PT polymerase chain reaction.
 XX
 PS Disclosure, SEQ ID NO 64; 419bp; English.
 XX
 CC The invention relates to a novel method for detecting adeno-associated
 CC virus (AAV) sequences in a sample, which comprises subjecting a sample
 CC containing a DNA to amplification via a polymerase chain reaction (PCR).
 CC The AAV sequence have the following activities: cytostatic,
 CC antidiabetic, antirheumatic, antiarthritic, neuroprotective,
 CC antipsoriatic, antithyroid, dermatological, and antiinflammatory. The AAV
 CC sequence can be used in gene therapy or as part of a vaccine to treat
 CC disorders. The method is useful in detecting and/or identifying AAV
 CC sequences and isolating novel sequences that are identified. The
 CC sequences may be used e.g. for preventing or treating hyperproliferative
 CC conditions such as cancers and psoriasis, and other autoimmune diseases
 CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
 CC thyroiditis, scleroderma or Crohn's disease. This sequence represents an
 CC AAV related protein sequence of the invention.
 XX
 SO Sequence 736 AA.

Query Match 100.0%; Score 3989; DB 7; Length 736;
 Best Local Similarity 100.0%; Pred. No. 5, 2e-311;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLENDLSGIREMMDLKGPAPKPKANOKODDGRGLVPGKYLGPFNGLD 60
 DB 1 MAADGYLPDWLENDLSGIREMMDLKGPAPKPKANOKODDGRGLVPGKYLGPFNGLD 60
 QY 61 KGEPPVNAADAALAEHDKAYDOOLKAGNPPYLRYNHADAERLOEDTSFGNIGRAVFO 120
 DB 61 KGEPPVNAADAALAEHDKAYDOOLKAGNPPYLRYNHADAERLOEDTSFGNIGRAVFO 120
 QY 121 AKKRVLEPLGLVEGATAPGKKRPVEQSPQEPDSSGIGKTGOAPAKRLNFGQTSDE 180
 DB 121 AKKRVLEPLGLVEGATAPGKKRPVEQSPQEPDSSGIGKTGOAPAKRLNFGQTSDE 180
 QY 121 AKKRVLEPLGLVEGATAPGKKRPVEQSPQEPDSSGIGKTGOAPAKRLNFGQTSDE 180
 DB 121 AKKRVLEPLGLVEGATAPGKKRPVEQSPQEPDSSGIGKTGOAPAKRLNFGQTSDE 180
 QY 181 SVDPPOPLGEBPPATPAVGPPTMASGGGAPWADNNEGADGVGNASGMWCHDSTWLGDRV 240
 DB 181 SVDPPOPLGEBPPATPAVGPPTMASGGGAPWADNNEGADGVGNASGMWCHDSTWLGDRV 240
 QY 241 TTSRTTAALPTVNNHLTKOISSASTGASNDNHYFGYSTPMGYPFENFPHCFSPRDMORL 300
 DB 241 TTSRTTAALPTVNNHLTKOISSASTGASNDNHYFGYSTPMGYPFENFPHCFSPRDMORL 300
 QY 241 TTSRTTAALPTVNNHLTKOISSASTGASNDNHYFGYSTPMGYPFENFPHCFSPRDMORL 300
 DB 241 TTSRTTAALPTVNNHLTKOISSASTGASNDNHYFGYSTPMGYPFENFPHCFSPRDMORL 300
 QY 301 INNMGFRPKRLNFKLFNIQYKEVTINDGVTTIANNLSTVQVSDSEYQLPYVLGSAHQ 360
 DB 301 INNMGFRPKRLNFKLFNIQYKEVTINDGVTTIANNLSTVQVSDSEYQLPYVLGSAHQ 360
 QY 361 GCLPPPADVPMIPQYGYLTILNNGSQAVGRSSPYCLEYFSPQMLRTGNNPFTSTPEEVP 420

DB 361 GCLPPPADVPMIPQYGYLTILNNGSQAVGRSSPYCLEYFSPQMLRTGNNPFTSTPEEVP 420
 QY 421 FHSSVYHQSGLDRLANPLIDQYLYLNKRTONGSGAQNODLFRSSPAGMSYQPKNWLP 480
 DB 421 FHSSVYHQSGLDRLANPLIDQYLYLNKRTONGSGAQNODLFRSSPAGMSYQPKNWLP 480
 QY 481 GPCYRQORYSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTMAHKKDEDEKFFPMGCV 540
 DB 481 GPCYRQORYSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTMAHKKDEDEKFFPMGCV 540
 QY 541 MIFGKESAGASNTALDNWMITDEEIKATNPATERFGTVAVNFQSSSTDPAQDVHAMG 600
 DB 541 MIFGKESAGASNTALDNWMITDEEIKATNPATERFGTVAVNFQSSSTDPAQDVHAMG 600
 QY 601 ALPGMTWODRDVYLQGPYIAKIPHTDGHFHPSPFMGFGKXKPPQILIKNTFVPANPPA 660
 DB 601 ALPGMTWODRDVYLQGPYIAKIPHTDGHFHPSPFMGFGKXKPPQILIKNTFVPANPPA 660
 QY 661 EFSATKFASTIYQSTGOVSEIEMELQKENSKRANPEVOYTSNYSKANVDFTVNNGL 720
 DB 661 EFSATKFASTIYQSTGOVSEIEMELQKENSKRANPEVOYTSNYSKANVDFTVNNGL 720
 QY 721 YTEPRPIGTRYLTRPL 736
 DB 721 YTEPRPIGTRYLTRPL 736

RESULT 5
 ADV70294
 ID ADV70294 standard; protein; 736 AA.
 XX
 AC ADV70294;
 XX
 DT 10-MAR-2005 (first entry)
 XX
 DE Primate adeno-associated virus 1 capsid protein VP1.
 XX
 KW immunosuppressive; gene therapy; immunogenicity; virus inactivation;
 KW hemophilia; peptic deficiency; galactosemia; phenylketonuria;
 KW Crigler-Najjar syndrome; Gout; Zellweger syndrome; ischemia;
 KW atherosclerosis; thrombosis; embolism; Parkinson's disease;
 KW congestive heart failure; cancer; inflammation; immune disorder;
 KW muscular dystrophy; diabetes; VP1.
 XX
 OS Adeno-associated virus 1.
 XX
 PN WO2004112727-A2.
 XX
 PD 29-DEC-2004.
 XX
 PF 21-JUN-2004; 2004WO-US019884.
 XX
 PR 19-JUN-2003; 2003US-0480395P.
 PR 30-APR-2004; 2004US-0567310P.
 PR 03-JUN-2004; 2004US-0576501P.
 XX
 PA (AVTG-) AVIGEN INC.
 XX
 PI Abetman AB, Colosi P, Lochrie MA, Surosky RT;
 XX
 DR WPI; 2005-048755/05.
 XX
 PT New mutated adeno-associated virus (AAV) capsid protein that when present
 PT in an AAV virion imparts decreased immunoreactivity to the virion as
 PT compared to the corresponding wild-type virion, useful for treating e.g.
 PT hemophilia.
 XX
 PS Example 5; SEQ ID NO 20; 136bp; English.
 XX
 CC The invention describes a mutated adeno-associated virus (AAV) capsid
 CC protein that when present in an AAV virion imparts decreased
 CC immunoreactivity to the virion as compared to the corresponding wild-type

CC virion. Also described are: a polynucleotide encoding the mutated protein above;
CC above; a recombinant AAV virion comprising the mutated protein above, and
CC delivering a recombinant AAV virion to a cell or tissue of a vertebrate
CC subject. The recombinant AAV virion is useful for delivering a
CC heterologous nucleic acid molecule to a cell or tissue of a vertebrate
CC subject, where the protein encoded by the heterologous nucleic acid
CC molecule is expressed at a level that provides a therapeutic effect.
CC where the recombinant AAV virion may comprise a non-primate, mammalian
CC AAV capsid protein that when present in an AAV virion imparts decreased
CC immunoreactivity to the virion as compared to immunoreactivity of primate
CC AAV-2, and the heterologous nucleic acid molecule, where the heterologous
CC nucleic acid molecule encodes a therapeutic protein and is operably
CC linked to control elements capable of directing the *in vivo* transcription
CC and translation of the protein. The protein or the recombinant AAV virion
CC is useful for treating or preventing a wide variety of disorders such as
CC hemophilia, glycerol storage deficiency type 1a, Papck deficiency,
CC galactosemia, phenylketonuria, Crigler-Najjar disease, Gout and Lesch-
CC Nyhan syndrome, Zellweger syndrome, ischemic diseases, atherosclerosis,
CC thrombosis, embolisms, Parkinson's disease, congestive heart failure,
CC cancer, inflammatory and immune disorders, muscular dystrophies, and
CC diabetes. This is the amino acid sequence of adeno-associated virus 1
CC (AAV1) capsid protein VP1.

SQ Sequence 736 AA;

Query Match 100.0%; Score 3989; DB 9; Length 736;

Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAAGVYLPEDMLIEDNLSEGI	REWMWDLKPGAPKPKAKQOQKDDRGVL	LPQYKYLGPENGLD	60	
Db	1	MAADGYLPDMLIEDNLSEGI	REWMWDLKPGAPKPKAKQOQKDDRGVL	LPQYKYLGPENGLD	60	
Qy	61	KGEVNNADAALAEHDKA	YDQOLKAGDMPYLALYNNHADADEFORL	DEPTSFSGNLGRAVRO	120	
Db	61	KGEVNNADAALAEHDKA	YDQOLKAGDMPYLALYNNHADADEFORL	DEPTSFSGNLGRAVRO	120	
Qy	121	AKKVELEBTLGVEEGAKTAPG	KPKRPAVEOSPQEPDSSSGIGTKGOQEP	AKKRLNFPQGTGDEE	180	
Db	121	AKKVELEBTLGVEEGAKTAPG	KPKRPAVEOSPQEPDSSSGIGTKGOQEP	AKKRLNFPQGTGDEE	180	
Qy	181	SVBPPOPILGEPPATPAAVG	PTTMAAGCGAPMADNNEGADUGNA	SGNNHICDSTWLGDRIT	240	
Db	181	SVBPPOPILGEPPATPAAVG	PTTMAAGCGAPMADNNEGADUGNA	SGNNHICDSTWLGDRIT	240	
Qy	241	TTSTRYMALPYNNHLYLKQI	SSASATGASDNHNYFGYSTPWG	FEDFNRFCHFSPRDMOYL	300	
Db	241	TTSTRYMALPYNNHLYLKQI	SSASATGASDNHNYFGYSTPWG	FEDFNRFCHFSPRDMOYL	300	
Qy	301	INNNMGPPRKLNFLFNIO	QKEVYNTTNOCVTTIANNLSTVO	VFSISEQLPVLGASAO	360	
Db	301	INNNMGPPRKLNFLFNIO	QKEVYNTTNOCVTTIANNLSTVO	VFSISEQLPVLGASAO	360	
Qy	361	GCLPPFPADVEMIPOYGYL	TLLNGSOAVGRSSFCLYEPPSO	MLRGNFTFSYTEEYP	420	
Db	361	GCLPPFPADVEMIPOYGYL	TLLNGSOAVGRSSFCLYEPPSO	MLRGNFTFSYTEEYP	420	
Qy	421	FHSSYASHOSLDRLNPL	IDOLYLYLNTTNOSSGANO	KDLFSGRSSPAGMSVOPKNMLP	480	
Db	421	FHSSYASHOSLDRLNPL	IDOLYLYLNTTNOSSGANO	KDLFSGRSSPAGMSVOPKNMLP	480	
Qy	481	GPCYRQORVSKTKTDNN	NNSNFTWTGASKYKYNLNGRES	IINPGTAMASHKODEKFPMSGV	540	
Db	481	GPCYRQORVSKTKTDNN	NNSNFTWTGASKYKYNLNGRES	IINPGTAMASHKODEKFPMSGV	540	
Qy	541	MIPKESAGANTLADNM	MITDEEELKATNPATATER	FGVVAANFPOSSSTDPATGVHANG	600	
Db	541	MIPKESAGANTLADNM	MITDEEELKATNPATATER	FGVVAANFPOSSSTDPATGVHANG	600	
Qy	601	ALPGWMDRODRVYL	OGPIWAKIPHTDGHF	HPSLPMLGFGKLP	POIILIKNTVPANPPA	660
Db	601	ALPGWMDRODRVYL	OGPIWAKIPHTDGHF	HPSLPMLGFGKLP	POIILIKNTVPANPPA	660

Qy	661	EFSATKPSFLLQYSTGQSVSEIEMELQENSKRNPEVQYTSNPAKSANVPFTYDNGI	720
Db	661	EFSATKPSFLLQYSTGQSVSEIEMELQENSKRNPEVQYTSNPAKSANVPFTYDNGI	720
Qy	721	YTEPPPIGTRYLTRPL	736
Db	721	YTEPPPIGTRYLTRPL	736

RESULT 6
AD227069 standard; protein, 736 AA.
AD227069

RESULT 6

ID ADZ27069 standard; protein; 736 AA.

AC ADZ27069;

DT 30-JUN-2005 (first entry)

DE Adeno-associated virus protein SEQ ID NO 219

KW rheumatoid arthritis; multiple sclerosis; sarcoidosis.

KW bacterial infection; cancer; ulcerative colitis; antirheumatic;

KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;

322

XX

XX

2000

XX

PR 29-APR-2004; 2004US-0566546P.

PA (TYPE-) UNIV PENNSYLVANIA.

PI Wilson JM, Gao G, Alvira MR, Vandenbergh LH;

DR WPI; 2005-285437/29.

PT New adeno-associated virus (AAV) clade comprising at least three AAV

PT, scleroderma, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial

XX

XX

CC at least three AAV members, where each member of the AAV clade is

by a bootstrap value of at least 75 % per 1000 isolates and a Poisson

compositions of the present invention are useful for the prevention

CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's disease, hepatitis, HIV, bacterial infection, cancer and ulcerative colitis.

CC collitis. The present sequence represents the amino acid sequence of an

X

Answer: Notch 100 A#, 66666 30000, DB 0, 10000 736.

Best Local Similarity 100.0%; Pred. NO. 5.2e-311;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps

Q. 1 MANDCVI.BNWI.FDNI.SEGTPEWMDI.KBCAPKPKANOKODNCPGI.YI.PGVKYI.GPENGI.D.6

1 МАДСЫ.РНТ.ЕДНТ.СЕРТ.ЕВМДТ.КРГА.РРКАНО.КОД.ГРГТ.УТ.РГК.УТ.ГРЕНГТ.Д. 6

61 KGEVNAADAAT.EHDKAYDOOT.KAGNUPYI.PYNHADAEEQERI.OEDTSEGGNI.GRAVEO 1

SQ Sequence 736 AA;

Query Match	100.0%;	Score 3989;	DB 9;	Length 736;
-------------	---------	-------------	-------	-------------

Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAGYLPDMLEDDNLSEGI REMWMDLPPKAPRKAAQOQODDPRGLVLPCKYKILGPPNGLD 60

QY 1 MAAAGYLPDMLEDDNLSEGI REMWMDLPPKAPRKAAQOQODDPRGLVLPCKYKILGPPNGLD 60

QY 1 MAAAGYLPDMLEDDNLSEGI REMWMDLPPKAPRKAAQOQODDPRGLVLPCKYKILGPPNGLD 60

QY 61 KGEPVNADAALAEHDKAYDQOLKAGDDPYLYRNHADAEFQERLOEDTSGFGNIGRAVTR 120

```

Db      |||||
61      KGEPVNAADAAALEHKAAYDQOLKAGNPLRLRYNHADAERLOEBTSGGNGRAVVFQ 120
Qy      121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTQOQPAKKRLNFGQTGDE 180
Db      121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTQOQPAKKRLNFGQTGDE 180
Qy      181 SVDPQPLGEPPTPAVGFPTTMAAGGAPMADNNEGADGVNAAGNWHCDSTWLGDRVI 240
Db      181 SVDPQPLGEPPTPAVGFPTTMAAGGAPMADNNEGADGVNAAGNWHCDSTWLGDRVI 240
Qy      241 TTSTRTVALPTYNHLYKOISSASTGASNDNHFGYSTPMGYDPFNRFHCFSPRDMQRL 300
Db      241 TTSTRTVALPTYNHLYKOISSASTGASNDNHFGYSTPMGYDPFNRFHCFSPRDMQRL 300
Qy      301 INNMGFRPKRLNFKLFNIQVKEVTNDGVTIANNLITSTVOYFSDSEYOLPYVLSAHQ 360
Db      301 INNMGFRPKRLNFKLFNIQVKEVTNDGVTIANNLITSTVOYFSDSEYOLPYVLSAHQ 360
Qy      361 GCLPPPADVFMIPQGYLTLLNNGSQAVGSSFCLEFPSSQMLRTGNNTFSTYTEEVP 420
Db      361 GCLPPPADVFMIPQGYLTLLNNGSQAVGSSFCLEFPSSQMLRTGNNTFSTYTEEVP 420
Qy      421 FHSSYAHOSGLDRLMPLIDQYLYLNRTONOGSAGNKKLLFSRGSPPAGMSVQPKWMLP 480
Db      421 FHSSYAHOSGLDRLMPLIDQYLYLNRTONOGSAGNKKLLFSRGSPPAGMSVQPKWMLP 480
Qy      481 GPCYRQORVSKTPTDNNNSNFTWTGASKYMLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
Db      481 GPCYRQORVSKTPTDNNNSNFTWTGASKYMLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
Qy      541 MIFGESAGASNTALDNVMTTDEBEIKATNPVATERGTVAANFQSSSTDPATGDVHAMG 600
Db      541 MIFGESAGASNTALDNVMTTDEBEIKATNPVATERGTVAANFQSSSTDPATGDVHAMG 600
Qy      601 ALGGMWODRDVYLOGIYAKIPHTDGHFHPSPIMGFGLKNPPOLLKNTTPVAPAPPA 660
Db      601 ALGGMWODRDVYLOGIYAKIPHTDGHFHPSPIMGFGLKNPPOLLKNTTPVAPAPPA 660
Qy      661 EFSATKFAPIITQYSTQGVSEIEMWELQKENSCKRMNBEVOYTSNYAKSANVDFTVDNGL 720
Db      661 EFSATKFAPIITQYSTQGVSEIEMWELQKENSCKRMNBEVOYTSNYAKSANVDFTVDNGL 720
Qy      721 YTEPRPIGTTRYLTRPL 736
Db      721 YTEPRPIGTTRYLTRPL 736

```

RESULT 7

```

AD227012
ID      AD227012 standard; protein; 736 AA.
XX
AC      AD227012;
XX
DT      30-JUN-2005 (first entry)
XX
DE      Adeno-associated virus protein SEQ ID NO 162.
XX
KW      rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW      scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW      bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW      anticholinergic; neuroprotective; antiinflammatory; antidiabetic;
KW      antipsoriatic; vasculitic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW      virucide; antibacterial; cytostatic; antineoplastic; dermatological.
XX
OS      Adeno-associated virus.
XX
PN      WO2005033321-A2.
XX
PD      14-APR-2005.
XX
PF      30-SEP-2004; 2004WO-US028817.
XX

```

```

PR      30-SEP-2003; 2003US-0508226P.
PR      29-APR-2004; 2004US-0566546P.
XX
XX      (TYPE-) UNIV PENNSYLVANIA.
XX
PI      Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX
XX      WPI; 2005-285437/29.
XX
XX      New adeno-associated virus (AAV) clade comprising at least three AAV
XX      members, useful for preventing and/or treating arthritis, multiple
XX      sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX      infection and cancer.
XX
XX      Disclosure, SEQ ID NO 162; 569pp; English.
XX
XX      The invention relates to an adeno-associated virus (AAV) clade comprising
XX      at least three AAV members, where each member of the AAV clade is
XX      phylogenetically related as determined using a Neighbor-joining heuristic
XX      by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX      correction distance measurement of no more than 0.05. The methods and
XX      compositions of the present invention are useful for the prevention
XX      and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX      Crohn's disease, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX      disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX      colitis. The present sequence represents the amino acid sequence of an
XX      adeno-associated virus protein.
XX
XX      Sequence 736 AA:
XX
XX      Query Match          100.0%; Score 3989; DB 9; Length 736;
XX      Best Local Similarity 100.0%; Pred. No. 5.2e-311;
XX      Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


QY 601 ALPGMWQDRDVLQGIWAKIPHTDGHFHPSPLMGSGFGLKNPPQILLIKNTVPANPPA 660
DB 601 ALPGMWQDRDVLQGIWAKIPHTDGHFHPSPLMGSGFGLKNPPQILLIKNTVPANPPA 660
QY 661 EFSATKFASTITQYSTQGVSEIEMELQKNSKRMNEVOYTSNVAKSANVDFTVNNGL 720
DB 661 EFSATKFASTITQYSTQGVSEIEMELQKNSKRMNEVOYTSNVAKSANVDFTVNNGL 720
QY 721 YTEPRPIGTRYLTRPL 736
DB 721 YTEPRPIGTRYLTRPL 736
RESULT 8
ADV67506
ID ADV67506 standard; protein; 735 AA.
XX
AC ADV67506;
XX
DT 10-MAR-2005 (first entry)
XX
DE Amino acid sequence of the capsid protein of AAV serotype 1.
XX
KM antiarteriosclerotic; antilipemic; gene therapy; cholesterol;
KM apolipoprotein E; apoe; apolipoprotein A; apoa; atherosclerosis;
KM lipoprotein defect; capsid protein.
XX
OS Adeno-associated virus.
XX
PN WO2004108922-A2.
XX
PD 16-DEC-2004.
XX
PF 23-APR-2004; 2004WO-US010965.
XX
PR 25-APR-2003; 2003JUS-0465293P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Rader DJ, Wilson JM;
XX
DR WPI; 2005-031700/03.
XX
PT Lowering total cholesterol levels and treating atherosclerosis in a
PT subject comprises delivering a recombinant adeno-associated virus (AAV)
PT comprising an AAV serotype capsid protein or a gene encoding human
PT apolipoprotein E (apoe) or apoa.
XX
PS Disclosure; SEQ ID NO 4; 69pp; English.
XX
CC The specification describes a method for lowering total cholesterol
CC levels in a subject. The method comprises delivering to the subject a
CC recombinant adeno-associated virus (AAV) comprising a gene encoding a
CC human apolipoprotein E (apoe) or apoa under the control of a regulatory
CC control sequence which directs expression of the gene. The recombinant
CC AAV also comprises a capsid protein selected from an AAV serotype (e.g.
CC AAV7 or AAV8) which preferentially expresses high levels of transgene in
CC live. A therapeutically effective amount of apoe or apoa expression is
CC obtained upon delivery of low dose of AAV. The method of the invention is
CC useful for lowering total cholesterol levels in a subject, e.g. for
CC treating atherosclerosis, and for correcting defects in lipoprotein. The
CC present sequence represents a capsid protein of AAV serotype 1 (AAV1),
CC which may be used in recombinant AAV vectors of the invention.
XX
SQ Sequence 735 AA;
Query Match 99.9%; Score 3985; DB 9; Length 735;
Best Local Similarity 100.0%; Pred. No. 1.1e-310;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAADGYLPDLIEDNLSCIREWMDLKGAEPKPAKNOCKODDGRGLVLPGYKYLGFPMGLD 60
|||||

DB 1 MAADGYLPDLIEDNLSCIREWMDLKGAEPKPAKNOCKODDGRGLVLPGYKYLGFPMGLD 60
QY 61 KGEFVNADAALAEHDKAYDOOLKAGDNPYLRYNHADAEPERLQEDTSPFGNIGRAVFQ 120
DB 61 KGEFVNADAALAEHDKAYDOOLKAGDNPYLRYNHADAEPERLQEDTSPFGNIGRAVFQ 120
QY 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPSSSGIGTGGQPAKKRLNFGQTGDS 180
DB 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPSSSGIGTGGQPAKKRLNFGQTGDS 180
QY 181 SVDPPOPPLGEPPTPAVGPPTTMASGGGAPMADNNEGADGVGNAGMWHCDSTLGDRI 240
DB 181 SVDPPOPPLGEPPTPAVGPPTTMASGGGAPMADNNEGADGVGNAGMWHCDSTLGDRI 240
QY 241 TTSRTWALPTYNHLYKQISASTGASNDNHYFGYSTPMGYPDNRFHCFSPRDQRL 300
DB 241 TTSRTWALPTYNHLYKQISASTGASNDNHYFGYSTPMGYPDNRFHCFSPRDQRL 300
QY 301 INNNMGFRPKRLNFKLFNIOVKEYTTNDGYTTTANNITSTYQVSDSEYQLPYLGSAAQ 360
DB 301 INNNMGFRPKRLNFKLFNIOVKEYTTNDGYTTTANNITSTYQVSDSEYQLPYLGSAAQ 360
QY 361 GCLPEPPADVPMIPOGYLTLLNGSQAVGSSPYCLEYPPSQMLRTGNNFTFSYTFEEVP 420
DB 361 GCLPEPPADVPMIPOGYLTLLNGSQAVGSSPYCLEYPPSQMLRTGNNFTFSYTFEEVP 420
QY 421 FHSYSAHSQSLDRIMNPLIDQVLYYLNRTONQSGSAQNKDLLFSRGSFAGMSVQPKWLP 480
DB 421 FHSYSAHSQSLDRIMNPLIDQVLYYLNRTONQSGSAQNKDLLFSRGSFAGMSVQPKWLP 480
QY 481 GPCYRQGRVSKTKTDNNNSNFTWTGASKYLNNGRESIINPGTAMASHKDEDEKFPMSGV 540
DB 481 GPCYRQGRVSKTKTDNNNSNFTWTGASKYLNNGRESIINPGTAMASHKDEDEKFPMSGV 540
QY 541 MIFGESAGASNTALDVMITDDEEIKATNPVATERGTAIVNFOSSSTDPAIDVAMG 600
DB 541 MIFGESAGASNTALDVMITDDEEIKATNPVATERGTAIVNFOSSSTDPAIDVAMG 600
QY 601 ALPGMWQDRDVLQGIWAKIPHTDGHFHPSPLMGSGFGLKNPPQILLIKNTVPANPPA 660
DB 601 ALPGMWQDRDVLQGIWAKIPHTDGHFHPSPLMGSGFGLKNPPQILLIKNTVPANPPA 660
QY 661 EFSATKFASTITQYSTQGVSEIEMELQKNSKRMNEVOYTSNVAKSANVDFTVNNGL 720
DB 661 EFSATKFASTITQYSTQGVSEIEMELQKNSKRMNEVOYTSNVAKSANVDFTVNNGL 720
QY 721 YTEPRPIGTRYLTRP 735
DB 721 YTEPRPIGTRYLTRP 735
RESULT 9
AD227007
ID AD227007 standard; protein; 736 AA.
XX
AC AD227007;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus protein SEQ ID NO 157.
XX
KM rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KM scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KM bacterial infection; cancer; ulcerative colitis; antirheumatic;
KM antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
KM antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KM virucide; antibacterial; cytostatic; antitumor; dermatological.
XX
OS Adeno-associated virus.
XX
PN WO2005033321-A2.
XX
PD 14-APR-2005.

XX 30-SEP-2004; 2004WO-US028817.
 PF 30-SEP-2003; 2003US-0508226P.
 XX 29-APR-2004; 2004US-0566546P.
 XX (UYPE-) UNIV PENNSYLVANIA.
 XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
 PI WPI; 2005-285437/29.
 DR
 XX New adeno-associated virus (AAV) clade comprising at least three AAV
 PT members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.
 XX
 PS Claim 19; SEQ ID NO 157; 569pp; English.
 XX
 CC The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is
 CC phylogenetically related as determined using a Neighbor-Joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents the amino acid sequence of an
 CC adeno-associated virus protein.
 CC
 XX
 SQ Sequence 736 AA;
 Query Match 99.3%; Score 3968; DB 9; Length 736;
 Best Local Similarity 99.3%; Pred. No. 2.5e-309;
 Matches 731; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLENDLSEGIHEWMDLKGAPKPRANQKODDGGGLVLPGYKYLGPFGGLD 60
 DB 1 MAADGYLPDWLENDLSEGIHEWMDLKGAPKPRANQKODDGGGLVLPGYKYLGPFGGLD 60
 QY 61 KGEPVNAADAALAEHDKAYDQOLKAGNPLYLRNHADEFOERLOEPTSGNLGRAVFO 120
 DB 61 KGEPVNAADAALAEHDKAYDQOLKAGNPLYLRNHADEFOERLOEPTSGNLGRAVFO 120
 QY 121 AKKRVLEPLGLVEGAKTAPGKRPVQSPQEPDSSSGIGKTQQPAKRLNFGQTGDS 180
 DB 121 AKKRVLEPLGLVEGAKTAPGKRPVQSPQEPDSSSGIGKTQQPAKRLNFGQTGDS 180
 QY 181 SVDPDQPLGPRPATPAVGTPTMASGGGAPMADNNEGADGVGNASGWHCDSTMLGRVI 240
 DB 181 SVDPDQPLGPRPATPAVGTPTMASGGGAPMADNNEGADGVGNASGWHCDSTMLGRVI 240
 QY 241 TTSRTTAAALPTYNHLYKQISSASTGASNDNHYGYSPTMGYPDFNRFCHFSRDMQRL 300
 DB 241 TTSRTTAAALPTYNHLYKQISSASTGASNDNHYGYSPTMGYPDFNRFCHFSRDMQRL 300
 QY 301 INNMGFRPRKLNFKLNIQVEVTTNDGVTTIANNLSTVQVSEDEYQLPYVLSAQ 360
 DB 301 INNMGFRPRKLNFKLNIQVEVTTNDGVTTIANNLSTVQVSEDEYQLPYVLSAQ 360
 QY 361 GCLPPPPADVFMIPQYGYTLNLNGSQAVGRSSFFCLEYFSPOMLRNTNNFTFSYTEBVP 420
 DB 361 GCLPPPPADVFMIPQYGYTLNLNGSQAVGRSSFFCLEYFSPOMLRNTNNFTFSYTEBVP 420
 QY 421 FHSSYAHOSLDRLMNPFLIDQVLYLNRTONOSGSAONKLLFSGSPAGMSVQPKWLP 480
 DB 421 FHSSYAHOSLDRLMNPFLIDQVLYLNRTONOSGSAONKLLFSGSPAGMSVQPKWLP 480
 QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEKFFPMGV 540
 DB 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEKFFPMGV 540

QY 541 MIFGKESAGASNTALNNMTDEEIKATNPVATERFGVAVNFQSSSTDTPATGDVHANG 600
 DB 541 MIFGKESAGASNTALNNMTDEEIKATNPVATERFGVAVNFQSSSTDTPATGDVHANG 600
 QY 601 ALPGMWQODRDVYLQGPIMAKIPHTDGHFHPSPLMGFGKXNPPQILIKNTEVPANPPA 660
 DB 601 ALPGMWQODRDVYLQGPIMAKIPHTDGHFHPSPLMGFGKXNPPQILIKNTEVPANPPA 660
 QY 661 EFSATKPFASFTIQYSTGVQSVLEIEMELOKENSKRKNPEVOYTSNYSKSNVDTVDNGL 720
 DB 661 EFSATKPFASFTIQYSTGVQSVLEIEMELOKENSKRKNPEVOYTSNYSKSNVDTVDNGL 720
 QY 721 YTEPRPIGTRVLYLRPL 736
 DB 721 YTEPRPIGTRVLYLRPL 736

RESULT 10
 AAB59847
 ID AAB59847 standard; protein; 736 AA.
 XX
 AC AAB59847;
 XX
 DT 28-MAR-2001 (first entry)
 XX
 DE AAV6 capsid protein VP1.
 XX
 KM AAV6; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
 KM atherosclerosis; sickle cell anemia; thalassemia;
 KM blood clotting disorder; diabetes; capsid protein VP1.
 XX
 OS Adeno associated virus.
 XX
 PN US6156303-A.
 XX
 PD 05-DEC-2000.
 XX
 PF 11-JUN-1997; 97US-00873168.
 XX
 PR 11-JUN-1997; 97US-00873168.
 XX
 PA (UNIW) UNIV WASHINGTON.
 XX
 PI Russell DW, Rutledge EA;
 XX
 DR WPI; 2001-060164/07.
 XX
 PT Adeno-associated virus serotype 6 and viral vector derived from it for
 PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
 PT syndrome, sickle cell anemia, thalassemia and diabetes.
 XX
 PS Disclosure; Fig 2; 50pp; English.
 XX
 CC The present invention relates to adeno-associated virus serotypes. The
 CC present sequence is capsid protein VP1 of one such serotype (AAV6). AAV6
 CC can be used to construct AAV viral vectors for use in gene therapy for a
 CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
 CC sickle cell anemia, thalassemia, blood clotting disorders and diabetes.
 CC The AAV viral vectors have increased transduction efficiency of a
 CC particular host cell as the AAV virion containing the AAV vector genome
 CC can be modified to express a capsid protein of an AAV serotype that
 CC transduces the selected host cell
 CC
 XX
 SQ Sequence 736 AA;
 Query Match 99.3%; Score 3963; DB 4; Length 736;
 Best Local Similarity 99.2%; Pred. No. 6.4e-309;
 Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLENDLSEGIHEWMDLKGAPKPRANQKODDGGGLVLPGYKYLGPFGGLD 60
 DB 1 MAADGYLPDWLENDLSEGIHEWMDLKGAPKPRANQKODDGGGLVLPGYKYLGPFGGLD 60


```

Qy 61 KGEVYNADAAALEHDKAYDOOLKAGDNPYLRYNHADAERFERLOEDTSFGGNIGRAVFO 120
Db 61 KGEVYNADAAALEHDKAYDOOLKAGDNPYLRYNHADAERFERLOEDTSFGGNIGRAVFO 120
Qy 121 AKKRVLEPGLIVEGAKTAPGKKRPVEOSPOEPDSSSGIGKTGQOPAKKRINFQOTGDS 180
Db 121 AKKRVLEPGLIVEGAKTAPGKKRPVEOSPOEPDSSSGIGKTGQOPAKKRINFQOTGDS 180
Qy 181 SVDPPOPLGEBPATPAAVGPTTMASSGGAPMADNNEGADGVGNASGMWCHDSTWLGDRVI 240
Db 181 SVDPPOPLGEBPATPAAVGPTTMASSGGAPMADNNEGADGVGNASGMWCHDSTWLGDRVI 240
Qy 241 TTSRTWALPTYNNHLKYQISSASTGASNDNHFGYSTPMGCFDPNRFCHFSRDMQRL 300
Db 241 TTSRTWALPTYNNHLKYQISSASTGASNDNHFGYSTPMGCFDPNRFCHFSRDMQRL 300
Qy 301 INNMGRPRKRLNFKLFNIOVKETVTDGVTIANNLSTVQVSDSEYQLPYVLSAHQ 360
Db 301 INNMGRPRKRLNFKLFNIOVKETVTDGVTIANNLSTVQVSDSEYQLPYVLSAHQ 360
Qy 361 GCLPPPADVFMIPQYGYLTLNNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYTFEBVP 420
Db 361 GCLPPPADVFMIPQYGYLTLNNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYTFEBVP 420
Qy 421 FHSSTYAHOSQSLDRLMNPLIDQYLYLNRTQNGSSAQNKDLLFSRGSFAGMSVOPKMWLP 480
Db 421 FHSSTYAHOSQSLDRLMNPLIDQYLYLNRTQNGSSAQNKDLLFSRGSFAGMSVOPKMWLP 480
Qy 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGV 540
Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGV 540
Qy 541 MIFGKSAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTPATGDVHAMG 600
Db 541 MIFGKSAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTPATGDVHAMG 600
Qy 601 ALRGMTVQDDVYIQGPIWAKIPHTDGHFHSPLMGGFGLKNPPQILIKTTPVPAHPA 660
Db 601 ALRGMTVQDDVYIQGPIWAKIPHTDGHFHSPLMGGFGLKNPPQILIKTTPVPAHPA 660
Qy 661 EFSATKASPIITQYSTQVSEIEMELOKENSKRMNPEVOYTSNVAASANDPFTVDNGL 720
Db 661 EFSATKASPIITQYSTQVSEIEMELOKENSKRMNPEVOYTSNVAASANDPFTVDNGL 720
Qy 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736

```

```

RESULT 11
ADE76566
ID ADE76566 standard; protein; 736 AA.
XX
AC ADE76566;
XX
DT 29-JAN-2004 (first entry)
XX
DE Adeno-associated virus (AAV) related protein, SEQ ID NO 65.
XX
KW adeno-associated virus; AAV; cytosolic; antipneumatic;
KW antitubercular; neuroprotective; antidiabetic; antithyroid;
KW dermatological; antiinflammatory; gene therapy; vaccine;
KW hyperproliferative; cancer; psoriasis; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; diabetes;
KW autoimmune thyroiditis; scleroderma; Crohn's disease.
XX
OS Unidentified.
XX
PN EP1310571-A2.
XX
PD 14-MAY-2003.
XX
PF 12-NOV-2002; 2002EP-00257826.

```

```

XX 13-NOV-2001; 2001US-0350607P.
PR 17-DEC-2001; 2001US-0341117P.
PR 01-MAY-2002; 2002US-0377066P.
PR 05-JUN-2002; 2002US-0386757P.
XX
PA (UNP-) UNIV PENNSYLVANIA.
XX
PI Gao G, Wilson JM, Alvira M;
XX
DR WPI; 2003-450984/43.
XX
PT Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT preventing or treating hyperproliferative or autoimmune diseases,
PT comprising subjecting a sample having a DNA to amplification via
PT polymerase chain reaction.
XX
PS Disclosure; SEQ ID NO 65; 419pp; English.
XX
CC The invention relates to a novel method for detecting adeno-associated
CC virus (AAV) sequences in a sample, which comprises subjecting a sample
CC containing a DNA to amplification via a polymerase chain reaction (PCR).
CC The AAV sequence have the following activities: cytostatic,
CC antipneumatic, antirheumatic, antitubercular, neuroprotective,
CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
CC sequence can be used in gene therapy or as part of a vaccine to treat
CC disorders. The method is useful in detecting and/or identifying AAV
CC sequences and isolating novel sequences that are identified. The
CC sequences may be used e.g. for preventing or treating hyperproliferative
CC conditions such as cancers and psoriasis, and other autoimmune diseases
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC thyroiditis, scleroderma or Crohn's disease. This sequence represents an
CC AAV related protein sequence of the invention.
XX
SQ Sequence 736 AA:

```

```

Query Match 99.3%; Score 3963; DB 7; Length 736;
Best Local Similarity 99.2%; Pred. No. 6.4e-309;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 MAADGYLPMLENDLSRGIREWMDLKPAPKPKANQKODDGRGLVPGYKYLGFENGLD 60
Db 1 MAADGYLPMLENDLSRGIREWMDLKPAPKPKANQKODDGRGLVPGYKYLGFENGLD 60
Qy 61 KGEVYNADAAALEHDKAYDOOLKAGDNPYLRYNHADAERFERLOEDTSFGGNIGRAVFO 120
Db 61 KGEVYNADAAALEHDKAYDOOLKAGDNPYLRYNHADAERFERLOEDTSFGGNIGRAVFO 120
Qy 121 AKKRVLEPGLIVEGAKTAPGKKRPVEOSPOEPDSSSGIGKTGQOPAKKRINFQOTGDS 180
Db 121 AKKRVLEPGLIVEGAKTAPGKKRPVEOSPOEPDSSSGIGKTGQOPAKKRINFQOTGDS 180
Qy 181 SVDPPOPLGEBPATPAAVGPTTMASSGGAPMADNNEGADGVGNASGMWCHDSTWLGDRVI 240
Db 181 SVDPPOPLGEBPATPAAVGPTTMASSGGAPMADNNEGADGVGNASGMWCHDSTWLGDRVI 240
Qy 241 TTSRTWALPTYNNHLKYQISSASTGASNDNHFGYSTPMGCFDPNRFCHFSRDMQRL 300
Db 241 TTSRTWALPTYNNHLKYQISSASTGASNDNHFGYSTPMGCFDPNRFCHFSRDMQRL 300
Qy 301 INNMGRPRKRLNFKLFNIOVKETVTDGVTIANNLSTVQVSDSEYQLPYVLSAHQ 360
Db 301 INNMGRPRKRLNFKLFNIOVKETVTDGVTIANNLSTVQVSDSEYQLPYVLSAHQ 360
Qy 361 GCLPPPADVFMIPQYGYLTLNNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYTFEBVP 420
Db 361 GCLPPPADVFMIPQYGYLTLNNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYTFEBVP 420
Qy 421 FHSSTYAHOSQSLDRLMNPLIDQYLYLNRTQNGSSAQNKDLLFSRGSFAGMSVOPKMWLP 480
Db 421 FHSSTYAHOSQSLDRLMNPLIDQYLYLNRTQNGSSAQNKDLLFSRGSFAGMSVOPKMWLP 480
Qy 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGV 540

```

```

DB      481 GPCYRQORVSKTTDNNNSFTWTGASKYVLRRESIINGGTAMASHKDDKDFPMSGV 540
QY      541 MIFGKSAGASNTALDNVMTTDEEBIKATNPVATERFGYAVANFOSSSTDPATGDVHAMG 600
DB      541 MIFGKSAGASNTALDNVMTTDEEBIKATNPVATERFGYAVANFOSSSTDPATGDVHAMG 600
QY      601 ALPGMWQORDVYVLOQPIWAKIPHTDGHFHPSPLMGFGFKHPPOILLKNTPTVPANPPA 660
DB      601 ALPGMWQORDVYVLOQPIWAKIPHTDGHFHPSPLMGFGFKHPPOILLKNTPTVPANPPA 660
QY      661 EFSATKFASTITQYSTGQVSVLEIEMELQKNSKRNKPEVOYTSNVAKSANVDFTVNNGL 720
DB      661 EFSATKFASTITQYSTGQVSVLEIEMELQKNSKRNKPEVOYTSNVAKSANVDFTVNNGL 720
QY      721 YTEPRPIGTRYLTRPL 736
DB      721 YTEPRPIGTRYLTRPL 736

RESULT 12
ADV70293
ID      ADV70293 standard; protein; 736 AA.
XX
AC      ADV70293;
XX
DT      10-MAR-2005 (first entry)
XX
DE      Primate adeno-associated virus 6 capsid protein VP1.
XX
KM      immunosuppressive; gene therapy; immunogenicity; virus inactivation;
KM      hemophilia; Pepck deficiency; galactosemia; phenylketonuria;
KM      Crigler-Najjar syndrome; Gout; Zellweger syndrome; ischemia;
KM      atherosclerosis; thrombosis; embolism; Parkinson's disease;
KM      congestive heart failure; cancer; inflammation; immune disorder;
KM      muscular dystrophy; diabetes; VPI.
XX
OS      Adeno-associated virus 6.
XX
PN      WO2004112727-A2.
XX
PD      29-DEC-2004.
XX
PF      21-JUN-2004; 2004MO-US019884.
XX
PR      19-JUN-2003; 2003US-0480395P.
PR      30-APR-2004; 2004US-0567310P.
PR      03-JUN-2004; 2004US-0576501P.
XX
PA      (AVIG-) AVIGEN INC.
PI      Arbetman AE, Colosi P, Lochrie MA, Surosky RT;
XX
DR      WPI; 2005-048755/05.
XX
PT      New mutated adeno-associated virus (AAV) capsid protein that when present
PT      in an AAV virion impacts decreased immunoreactivity to the virion as
PT      compared to the corresponding wild-type virion, useful for treating e.g.
PT      hemophilia.
XX
PS      Example 5; SEQ ID NO 19; 136pp; English.
XX
CC      The invention describes a mutated adeno-associated virus (AAV) capsid
CC      protein that when present in an AAV virion impacts decreased
CC      immunoreactivity to the virion as compared to the corresponding wild-type
CC      virion. Also described are: a polynucleotide encoding the mutated protein
CC      above; a recombinant AAV virion comprising the mutated protein above; and
CC      delivering a recombinant AAV virion to a cell or tissue of a vertebrate
CC      subject. The recombinant AAV virion is useful for delivering a
CC      heterologous nucleic acid molecule to a cell or tissue of a vertebrate
CC      subject, where the protein encoded by the heterologous nucleic acid
CC      molecule is expressed at a level that provides a therapeutic effect,
CC      where the recombinant AAV virion may comprise a non-primate, mammalian

```

```

CC      AAV capsid protein that when present in an AAV virion impacts decreased
CC      immunoreactivity to the virion as compared to immunoreactivity of primate
CC      AAV-2, and the heterologous nucleic acid molecule, where the heterologous
CC      nucleic acid molecule encodes a therapeutic protein and is operably
CC      linked to control elements capable of directing the in vivo transcription
CC      and translation of the protein. The protein or the recombinant AAV virion
CC      is useful for treating or preventing a wide variety of disorders such as
CC      hemophilia, glycogen storage deficiency type 1A, Pepck deficiency,
CC      galactosemia, phenylketonuria, Crigler-Najjar disease, Gout and Lesch-
CC      Nyan syndrome, Zellweger syndrome, ischemic diseases, atherosclerosis,
CC      thrombosis, embolisms, Parkinson's disease, congestive heart failure,
CC      cancer, inflammatory and immune disorders, muscular dystrophies, and
CC      diabetes. This is the amino acid sequence of adeno-associated virus 6
CC      (AAV6) capsid protein VP1.
XX
SQ      Sequence 736 AA;
XX
Query Match          99.3%; Score 3963; DB 9; Length 736;
Best Local Similarity 99.2%; Pred. No. 6,4e-309;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 MAADGYLPWLEDNLSSEGREWMDLKPAPKPAKNOQKDDGGLVLPQYKYLGPFGND 60
DB      1 MAADGYLPWLEDNLSSEGREWMDLKPAPKPAKNOQKDDGGLVLPQYKYLGPFGND 60
QY      61 KGPVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAERFOERLOEDTSFGNLAGRAVFQ 120
DB      61 KGPVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAERFOERLOEDTSFGNLAGRAVFQ 120
QY      121 AKKRVLPLGLVEEGAKTAPGKKRPVEQSPQEBSSSGIGTQGOAPAKRLNFGQTGSE 180
DB      121 AKKRVLPLGLVEEGAKTAPGKKRPVEQSPQEBSSSGIGTQGOAPAKRLNFGQTGSE 180
QY      181 SVDPDPLGEPATTPAAGPTTMASSGGAPMADNNGADGCGAASGNWHCDSTWLDGRI 240
DB      181 SVDPDPLGEPATTPAAGPTTMASSGGAPMADNNGADGCGAASGNWHCDSTWLDGRI 240
QY      241 TSTRTMAVPTYNHLYKQISSASTGASNDNHVFGYSTPAGYDFNRFHCFSPRDMQR 300
DB      241 TSTRTMAVPTYNHLYKQISSASTGASNDNHVFGYSTPAGYDFNRFHCFSPRDMQR 300
QY      301 INNNMGFRPKRLNFKLNIQVKEVTNDGVTTIANNLTSTVOVPSDEYQLPYVLSAQ 360
DB      301 INNNMGFRPKRLNFKLNIQVKEVTNDGVTTIANNLTSTVOVPSDEYQLPYVLSAQ 360
QY      361 GCLPPPADVFMIPQYGYLTNNGSQAVGRSSVYCLEYFSPQMLRTGNNFTFSYTEEDVP 420
DB      361 GCLPPPADVFMIPQYGYLTNNGSQAVGRSSVYCLEYFSPQMLRTGNNFTFSYTEEDVP 420
QY      421 FHSYVAHSQSLDKLMPPLIDQYLYLNRTONGSQAONKDLFSRGSPPAGMSVQPKWLP 480
DB      421 FHSYVAHSQSLDKLMPPLIDQYLYLNRTONGSQAONKDLFSRGSPPAGMSVQPKWLP 480
QY      481 GPCYRQORVSKTTDNNNSFTWTGASKYVLRRESIINGGTAMASHKDDKDFPMSGV 540
DB      481 GPCYRQORVSKTTDNNNSFTWTGASKYVLRRESIINGGTAMASHKDDKDFPMSGV 540
QY      541 MIFGKSAGASNTALDNVMTTDEEBIKATNPVATERFGYAVANFOSSSTDPATGDVHAMG 600
DB      541 MIFGKSAGASNTALDNVMTTDEEBIKATNPVATERFGYAVANFOSSSTDPATGDVHAMG 600
QY      601 ALPGMWQORDVYVLOQPIWAKIPHTDGHFHPSPLMGFGFKHPPOILLKNTPTVPANPPA 660
DB      601 ALPGMWQORDVYVLOQPIWAKIPHTDGHFHPSPLMGFGFKHPPOILLKNTPTVPANPPA 660
QY      661 EFSATKFASTITQYSTGQVSVLEIEMELQKNSKRNKPEVOYTSNVAKSANVDFTVNNGL 720
DB      661 EFSATKFASTITQYSTGQVSVLEIEMELQKNSKRNKPEVOYTSNVAKSANVDFTVNNGL 720
QY      721 YTEPRPIGTRYLTRPL 736
DB      721 YTEPRPIGTRYLTRPL 736

```

RESULT 13
AD227070
ID AD227070 standard; protein; 736 AA.
XX
AC AD227070;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus protein SEQ ID NO 220.
XX
KM rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KM scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KM bacterial infection; cancer; ulcerative colitis; antirheumatic;
KM antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
KM antidiabetic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KM virucide; antibacterial; cytostatic; antitumor; dermatological.
XX
OS Adeno-associated virus.
XX
PN WO2005033321-A2.
XX
PD 14-APR-2005.
XX
PF 30-SEP-2004; 2004MO-US028817.
XX
PR 30-SEP-2003; 2003US-0508226P.
XX
PR 29-APR-2004; 2004US-0566546P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX
DR WPI; 2005-285437/29.
XX
PT New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX
PS Disclosure; SEQ ID NO 220; 569pp; English.
XX
CC The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents the amino acid sequence of an
CC adeno-associated virus protein.
XX
SQ Sequence 736 AA;
XX
Query Match 99.3%; Score 3963; DB 9; Length 736;
Best Local Similarity 99.2%; Pred. No. 6,4e-309;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

DB 181 SVDPQPLGEPPTPAAVGTTTMASSGGAPMADNNEGADVGNAGMHCDSITWLGDRVI 240
QY 241 TTSTRTALPTVNNHLTKQISSASTGANDNHYFGVSTPGVDFNRFHCHFSBRDQRL 300
DB 241 TTSTRTALPTVNNHLTKQISSASTGANDNHYFGVSTPGVDFNRFHCHFSBRDQRL 300
QY 301 INNNWGRPRKLNPKLFNIOVKEVTNDGVTTIANNLTSIVQVFSDEYOLPYLGSAAHQ 360
DB 301 INNNWGRPRKLNPKLFNIOVKEVTNDGVTTIANNLTSIVQVFSDEYOLPYLGSAAHQ 360
QY 361 GCLPPPADVFMIPQVGYLTNNGSQAVGRSSFCLEYPFSOMRTGNFFSFTFEV 420
DB 361 GCLPPPADVFMIPQVGYLTNNGSQAVGRSSFCLEYPFSOMRTGNFFSFTFEV 420
QY 421 FHSYAHQSGLDLMLNPIDQYLYLNRTONOGSQAQNKDLFSRGSFAGMSVQPKWLP 480
DB 421 FHSYAHQSGLDLMLNPIDQYLYLNRTONOGSQAQNKDLFSRGSFAGMSVQPKWLP 480
QY 481 GPCYRQQRVSKTYTDNNNSNFTWTGASKYNLNGRESIINFGTAMASHKDEDFPMG 540
DB 481 GPCYRQQRVSKTYTDNNNSNFTWTGASKYNLNGRESIINFGTAMASHKDEDFPMG 540
QY 541 MIFKESAGASNTALDNWMTTDEEIKATNPVATERGTVAVNPOSSSTDPAQDVAMG 600
DB 541 MIFKESAGASNTALDNWMTTDEEIKATNPVATERGTVAVNPOSSSTDPAQDVAMG 600
QY 601 ALPGMWQDRDVLQGPIMAKIPHTDGHFPPSPLMGFGKPNPQILIKTPVPANPPA 660
DB 601 ALPGMWQDRDVLQGPIMAKIPHTDGHFPPSPLMGFGKPNPQILIKTPVPANPPA 660
QY 661 EFSATKFASTITQSTGVSEIEWELQKENSKRMEVEOYTSNYAKSANVDFTVNNGL 720
DB 661 EFSATKFASTITQSTGVSEIEWELQKENSKRMEVEOYTSNYAKSANVDFTVNNGL 720
QY 721 YTEPRPIGRTYLRPL 736
DB 721 YTEPRPIGRTYLRPL 736
RESULT 14
AD227086
ID AD227086 standard; protein; 737 AA.
XX
AC AD227086;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus protein SEQ ID NO 236.
XX
KM rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KM scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KM bacterial infection; cancer; ulcerative colitis; antirheumatic;
KM antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
KM antidiabetic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KM virucide; antibacterial; cytostatic; antitumor; dermatological.
XX
OS Adeno-associated virus.
XX
PN WO2005033321-A2.
XX
PD 14-APR-2005.
XX
PF 30-SEP-2004; 2004MO-US028817.
XX
PR 30-SEP-2003; 2003US-0508226P.
XX
PR 29-APR-2004; 2004US-0566546P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX
DR WPI; 2005-285437/29.

```
XX New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX
PS Claim 19; SEQ ID NO 236; 569pp; English.
XX
CC The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents the amino acid sequence of an
CC adeno-associated virus protein.
XX
XX Sequence 737 AA;
SQ
Query Match          97.7%; Score 3897.5; DB 9; Length 737;
Best Local Similarity 97.7%; Pred. No. 1.2e-303;
Matches 720; Conservative 6; Mismatches 10; Indels 1; Gaps 1;
QY 1 MAADGYLPWLENDLSEGIKREWMDLKPGAPKPRANQKODDGRGLVPGYKYLGPFGD 60
DB 1 MAADGYLPWLENDLSEGIKREWMDLKPGAPKPRANQKODDGRGLVPGYKYLGPFGD 60
QY 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAERQERLOEDTSGNLRGAVFQ 120
DB 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAERQERLOEDTSGNLRGAVFQ 120
QY 121 AKKRVLEPLGLVEGAKTAPGKRPVEQSPQ-EPDSSSGIGTKGQOPAKKRLNFGOTGDS 179
DB 121 AKKRVLEPLGLVEGAKTAPGKRPVEQSPQ-EPDSSSGIGTKGQOPAKKRLNFGOTGDS 180
QY 180 ESYVDPQPLGEPPTPAVAGPTTMAAGGAPMADNNEGADGVNAGSNHCHDSTWLGDRV 239
DB 180 ESYVDPQPLGEPPTPAVAGPTTMAAGGAPMADNNEGADGVNAGSNHCHDSTWLGDRV 240
QY 181 ESYVDPQPLGEPPTPAVAGPTTMAAGGAPMADNNEGADGVNAGSNHCHDSTWLGDRV 240
DB 181 ESYVDPQPLGEPPTPAVAGPTTMAAGGAPMADNNEGADGVNAGSNHCHDSTWLGDRV 240
QY 240 ITTSTETWALPTNNHLYKOISASGASNDNHYFGSTPWGVPEDPRFHCHSPDMDOR 239
DB 240 ITTSTETWALPTNNHLYKOISASGASNDNHYFGSTPWGVPEDPRFHCHSPDMDOR 239
QY 241 ITTSTETWALPTNNHLYKOISASGASNDNHYFGSTPWGVPEDPRFHCHSPDMDOR 300
DB 241 ITTSTETWALPTNNHLYKOISASGASNDNHYFGSTPWGVPEDPRFHCHSPDMDOR 300
QY 300 LNNNGGFRPKRLNFKLFNIQVKEVTINDGVTIANNLSTVQVFSDEQLPYVLSAH 359
DB 300 LNNNGGFRPKRLNFKLFNIQVKEVTINDGVTIANNLSTVQVFSDEQLPYVLSAH 359
QY 301 LNNNGGFRPKRLNFKLFNIQVKEVTINDGVTIANNLSTVQVFSDEQLPYVLSAH 360
DB 301 LNNNGGFRPKRLNFKLFNIQVKEVTINDGVTIANNLSTVQVFSDEQLPYVLSAH 360
QY 360 QGCLPEPPADVEMIPOYGYLTNNGSOAVGRSSPYCLEYFSPQMLRTGNFTFSYTFEYV 419
DB 360 QGCLPEPPADVEMIPOYGYLTNNGSOAVGRSSPYCLEYFSPQMLRTGNFTFSYTFEYV 420
QY 420 PFHSSVAHSQSLRLNPLLDQYLYLNTRONOGSAGNDLLFSRSPAGSMVQPPNWL 479
DB 420 PFHSSVAHSQSLRLNPLLDQYLYLNTRONOGSAGNDLLFSRSPAGSMVQPPNWL 480
QY 421 PLHSSVAHSQSLRLNPLLDQYLYLNTRONOGSAGNDLLFSRSPAGSMVQPPNWL 480
DB 421 PLHSSVAHSQSLRLNPLLDQYLYLNTRONOGSAGNDLLFSRSPAGSMVQPPNWL 480
QY 480 PGPCYRQORVSKTGTDDNNNSFTWTGASKYNLNGRESIIPGTAMASHKDDDEKFFPMG 539
DB 480 PGPCYRQORVSKTGTDDNNNSFTWTGASKYNLNGRESIIPGTAMASHKDDDEKFFPMG 540
QY 481 PGPCYRQORVSKTGTDDNNNSFTWTGASKYNLNGRESIIPGTAMASHKDDDEKFFPMG 540
DB 481 PGPCYRQORVSKTGTDDNNNSFTWTGASKYNLNGRESIIPGTAMASHKDDDEKFFPMG 540
QY 540 VMI FGESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAQDVHAM 600
DB 540 VMI FGESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAQDVHAM 600
QY 600 GALPGWMDORDVYLOGPIWAKIPTHGHPSPIMGFGGLKNPPDILKNPVPANPP 659
DB 600 GALPGWMDORDVYLOGPIWAKIPTHGHPSPIMGFGGLKNPPDILKNPVPANPP 660
QY 601 GALPGWMDORDVYLOGPIWAKIPTHGHPSPIMGFGGLKNPPDILKNPVPANPP 660
DB 601 GALPGWMDORDVYLOGPIWAKIPTHGHPSPIMGFGGLKNPPDILKNPVPANPP 660
QY 660 AEFSAATFASFITQYSTGVSVLEIWELOKENSKRNPVEQYTSNVAKSASVDPVTYNNG 719
DB 660 AEFSAATFASFITQYSTGVSVLEIWELOKENSKRNPVEQYTSNVAKSASVDPVTYNNG 720
```

```
QY 720 LYTEPRPICTRYLTRPL 736
DB 721 LYTEPRPICTRYLTRPL 737
RESULT 15
AD227010
ID AD227010 standard; protein; 737 AA.
XX
AC AD227010;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus protein SEQ ID NO 160.
XX
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antineumatic;
KW antirheumatic; neuroprotective; antiinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antiulcer; dermatological.
XX
OS Adeno-associated virus.
XX
PN MO2005033321-A2.
XX
PD 14-APR-2005.
XX
PF 30-SEP-2004; 2004MO-US028817.
XX
PR 30-SEP-2003; 2003US-0508226P.
PR 29-APR-2004; 2004US-0566546P.
XX
PA (UYPE-) UNTV PENNSYLVANIA.
XX
PI Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX
DR WPI; 2005-285437/29.
XX
PT New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX
PS Claim 46; SEQ ID NO 160; 569pp; English.
XX
CC The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents the amino acid sequence of an
CC adeno-associated virus protein.
XX
XX Sequence 737 AA;
SQ
Query Match          97.4%; Score 3883.5; DB 9; Length 737;
Best Local Similarity 97.4%; Pred. No. 1.0e-302;
Matches 718; Conservative 6; Mismatches 12; Indels 1; Gaps 1;
QY 1 MAADGYLPWLENDLSEGIKREWMDLKPGAPKPRANQKODDGRGLVPGYKYLGPFGD 60
DB 1 MAADGYLPWLENDLSEGIKREWMDLKPGAPKPRANQKODDGRGLVPGYKYLGPFGD 60
QY 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAERQERLOEDTSGNLRGAVFQ 120
DB 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAERQERLOEDTSGNLRGAVFQ 120
```

Qy	121	AKKVLIEPIGLYBEGAKTAPKKRPVJESQ - EPDSSGIGTKTGOOPAKKRLNFGQTGDS	179
Db	121	AKKVLIEPIGLYBEGAKTAPKKRPVJESQ - EPDSSGIGTKTGOOPAKKRLNFGQTGDS	180
Qy	180	ESVPPPOPIGEEPPATPAAGVPTTMASSGGGAPMADNNNEGADGVNAGSNNHCDSTWIGDRV	239
Db	181	ESVPPPOPIGEEPPATPAAGVPTTMASSGGGAPMADNNNEGADGVNAGSNNHCDSTWIGDRV	240
Qy	240	ITTSRTWALPTYNNHLYKQISSASGASNDNHFGYSTPMGYFDFNRFHCHSPRDMOR	239
Db	241	ITTSRTWALPTYNNHLYKQISSASGASNDNHFGYSTPMGYFDFNRFHCHSPRDMOR	300
Qy	300	LINNMWGPFRKLNFLFLENIOYKEVTTNDGVTIANNLSTVOYPSDSEYQULPYVGSAA	359
Db	301	LINNMWGPFRKLNFLFLENIOYKEVTTNDGVTIANNLSTVOYPSDSEYQULPYVGSAA	360
Qy	360	OGCLPPFPADVEMIPOYGLYLTNNGSQAVERSSFCLEYFPSSOMLRTGNNTFESYFEEV	419
Db	361	OGCLPPFPADVEMIPOYGLYLTNNGSQAVERSSFCLEYFPSSOMLRTGNNTFESYFEEV	420
Qy	420	PHSSYVASHQSIDLRLNPLIDOYLYLNTNONGSSAONKDLFSGSPAGMSVOPKANL	479
Db	421	PHSSYVASHQSIDLRLNPLIDOYLYLNTNONGSSAONKDLFSGSPAGMSVOPKANL	480
Qy	480	PGPCRRQOVSSTKTDNNNSNFWMTASAKYNNLGRSSINPGTAMASHODEKPFPMG	539
Db	481	PGPCRRQOVSSTKTDNNNSNFWMTASAKYNNLGRSSINPGTAMASHODEKPFPMG	540
Qy	540	VMI PGKESAGASNTALDNNMTDEEBEIKATNPVATERFGTVAVNFSSSTDPATGVHAM	599
Db	541	VMI PGKESAGASNTALDNNMTDEEBEIKATNPVATERFGTVAVNFSSSTDPATGVHAM	600
Qy	600	GALPGMWODROVYLOGPIWAKIPHTDGHFHPSPLMGGGGLKNPPOOLIKNTPVANPP	659
Db	601	GALPGMWODROVYLOGPIWAKIPHTDGHFHPSPLMGGGGLKNPPOOLIKNTPVANPP	660
Qy	660	AEFSAATKFASTFOYSTGVGSVEIWELEKENSKRNNPEVOYTSNPAKSANNVFTVDNNG	719
Db	661	AEFSAATKFASTFOYSTGVGSVEIWELEKENSKRNNPEVOYTSNPAKSANNVFTVDNNG	720
Qy	720	LYTEPRPIGTIRYLTAPL 736	
Db	721	LYTEPRPIGTIRYLTAPL 737	

THIS PAGE BLANK (USPTO,


```
Db      241  |||||||
Qy      301  INNMGRPRKLNFLKFNIOKEVTNDGVTIIANNLSTVQVFSDEYOLPVYLSAHQ 360
Db      301  INNMGRPRKLNFLKFNIOKEVTNDGVTIIANNLSTVQVFSDEYOLPVYLSAHQ 360
Qy      361  GCLPPPADVFMIPQGYLTLNNGSOAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEEVP 420
Db      361  GCLPPPADVFMIPQGYLTLNNGSOAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEEVP 420
Qy      421  FHSYVSHSOSLDLMLPDLIDQYLTYLNRTONOGSAONKDLFSRGSFAGMSYQPKWLP 480
Db      421  FHSYVSHSOSLDLMLPDLIDQYLTYLNRTONOGSAONKDLFSRGSFAGMSYQPKWLP 480
Qy      481  GPCYRQORVSKTIDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMG 540
Db      481  GPCYRQORVSKTIDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMG 540
Qy      541  MIFGKSAGASNTALDNVMTDEBEIKATNPVATERGTVAVNFQSSSTDPAIGDVHAMG 600
Db      541  MIFGKSAGASNTALDNVMTDEBEIKATNPVATERGTVAVNFQSSSTDPAIGDVHAMG 600
Qy      601  ALPGMWQODRDVYLQGFIMAKIPHTDGHFSPMLMGFGKKNPPOLLIKNTVPANPPA 660
Db      601  ALPGMWQODRDVYLQGFIMAKIPHTDGHFSPMLMGFGKKNPPOLLIKNTVPANPPA 660
Qy      661  EFSATKFAFIITQYSTQGVSEIEMWELQKENSKRMPNEVOYTSNVAKSANVDFVNDNGL 720
Db      661  EFSATKFAFIITQYSTQGVSEIEMWELQKENSKRMPNEVOYTSNVAKSANVDFVNDNGL 720
Qy      721  YTEPRPIGTYYLTRPL 736
Db      721  YTEPRPIGTYYLTRPL 736
```

```
RESULT 2
US-10-423-704A-5
; Sequence 5, Application US/10423704A
; Publication No. US20030228282A1
; GENERAL INFORMATION:
; APPLICANT: Geo. Guangdong
; APPLICANT: Wilson, James M.
; APPLICANT: Alviria, Mauricio
; TITLE OF INVENTION: Adeno-Associated Virus (AAV) Serotype 8 Sequences, Vectors
; FILE REFERENCE: UPN-02733AUSA
; CURRENT APPLICATION NUMBER: US/10/423, 704A
; PRIOR FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 60/341,151
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,123
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,122
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: PCT/US02/33630
; PRIOR FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 736
; TYPE: PRT
; ORGANISM: adeno-associated virus serotype 1
US-10-423-704A-5
```

```
Query Match      100.0%; Score 3989; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No.9.5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      61  KGEPVNAADAAAEHDKAVDQOLKAGDNPLYRNYHDAEFOERLOEDTSGNIGRAVFQ 120
Db      61  KGEPVNAADAAAEHDKAVDQOLKAGDNPLYRNYHDAEFOERLOEDTSGNIGRAVFQ 120
Qy      121  AKKRVLEPLGLVEGAKTAPGKRPEQSFQEPDSSSGIGTKTQOPAKKRLNFGQTGSE 180
Db      121  AKKRVLEPLGLVEGAKTAPGKRPEQSFQEPDSSSGIGTKTQOPAKKRLNFGQTGSE 180
Qy      181  SVDPDPLGEPAPTPAAGVPTTMASSGGAPMADNBNAGDGVNAGSGMWHCDSTWLDGRI 240
Db      181  SVDPDPLGEPAPTPAAGVPTTMASSGGAPMADNBNAGDGVNAGSGMWHCDSTWLDGRI 240
Qy      241  TSTRTMALPTNNHLYKQISSASTGASNDNHFGYSTPMGYFDENRFBCHFSRDMQRL 300
Db      241  TSTRTMALPTNNHLYKQISSASTGASNDNHFGYSTPMGYFDENRFBCHFSRDMQRL 300
Qy      301  INNMGRPRKLNFLKFNIOKEVTNDGVTIIANNLSTVQVFSDEYOLPVYLSAHQ 360
Db      301  INNMGRPRKLNFLKFNIOKEVTNDGVTIIANNLSTVQVFSDEYOLPVYLSAHQ 360
Qy      361  GCLPPPADVFMIPQGYLTLNNGSOAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEEVP 420
Db      361  GCLPPPADVFMIPQGYLTLNNGSOAVGRSSFYCLEYFPPSOMLRTGNFTFSYTFEEVP 420
Qy      421  FHSYVSHSOSLDLMLPDLIDQYLTYLNRTONOGSAONKDLFSRGSFAGMSYQPKWLP 480
Db      421  FHSYVSHSOSLDLMLPDLIDQYLTYLNRTONOGSAONKDLFSRGSFAGMSYQPKWLP 480
Qy      481  GPCYRQORVSKTIDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMG 540
Db      481  GPCYRQORVSKTIDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKODEDKFFPMG 540
Qy      541  MIFGKSAGASNTALDNVMTDEBEIKATNPVATERGTVAVNFQSSSTDPAIGDVHAMG 600
Db      541  MIFGKSAGASNTALDNVMTDEBEIKATNPVATERGTVAVNFQSSSTDPAIGDVHAMG 600
Qy      601  ALPGMWQODRDVYLQGFIMAKIPHTDGHFSPMLMGFGKKNPPOLLIKNTVPANPPA 660
Db      601  ALPGMWQODRDVYLQGFIMAKIPHTDGHFSPMLMGFGKKNPPOLLIKNTVPANPPA 660
Qy      661  EFSATKFAFIITQYSTQGVSEIEMWELQKENSKRMPNEVOYTSNVAKSANVDFVNDNGL 720
Db      661  EFSATKFAFIITQYSTQGVSEIEMWELQKENSKRMPNEVOYTSNVAKSANVDFVNDNGL 720
Qy      721  YTEPRPIGTYYLTRPL 736
Db      721  YTEPRPIGTYYLTRPL 736
```

```
RESULT 3
US-10-696-261-3
; Sequence 3, Application US/10696261
; Publication No. US20040057931A1
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/10/696,261
; PRIOR FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
```

US-10-696-261-3

Query Match 100.0%; Score 3989; DB 4; Length 736;
 Best Local Similarity 100.0%; Pred. No. 9.5e-309;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAGGYLPDMLEDMISGIRREMDLKGAPKPKANQKQODDGRGLVLPGYKYLGPFGNLD 60
 1 MAAGGYLPDMLEDMISGIRREMDLKGAPKPKANQKQODDGRGLVLPGYKYLGPFGNLD 60
 DB 1 MAAGGYLPDMLEDMISGIRREMDLKGAPKPKANQKQODDGRGLVLPGYKYLGPFGNLD 60
 QY 61 KGEVNAADAAALAEHDKAYDQOLKAGDNPYLRYNHADAERLQEDTSFGNLAGRAVQ 120
 61 KGEVNAADAAALAEHDKAYDQOLKAGDNPYLRYNHADAERLQEDTSFGNLAGRAVQ 120
 DB 61 KGEVNAADAAALAEHDKAYDQOLKAGDNPYLRYNHADAERLQEDTSFGNLAGRAVQ 120
 QY 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGGQPAKKRLNFGQTGSE 180
 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGGQPAKKRLNFGQTGSE 180
 DB 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGGQPAKKRLNFGQTGSE 180
 QY 181 SVDPQOLGEPPTPAVGPPTTMASSGGA PMADNNEGADGVGNASGNMHCSTWLGDRVI 240
 181 SVDPQOLGEPPTPAVGPPTTMASSGGA PMADNNEGADGVGNASGNMHCSTWLGDRVI 240
 DB 181 SVDPQOLGEPPTPAVGPPTTMASSGGA PMADNNEGADGVGNASGNMHCSTWLGDRVI 240
 QY 241 TTSRTMALPTNNHLKQISSASTGASNDNHFGYSTPMGYPDFNRFCHFSRDMQRL 300
 241 TTSRTMALPTNNHLKQISSASTGASNDNHFGYSTPMGYPDFNRFCHFSRDMQRL 300
 DB 241 TTSRTMALPTNNHLKQISSASTGASNDNHFGYSTPMGYPDFNRFCHFSRDMQRL 300
 QY 301 INNNWGRPKRLNFKLFNIQVEYTTNDGVTIANNLSTVQVSDSEYQLPYLGSANQ 360
 301 INNNWGRPKRLNFKLFNIQVEYTTNDGVTIANNLSTVQVSDSEYQLPYLGSANQ 360
 DB 301 INNNWGRPKRLNFKLFNIQVEYTTNDGVTIANNLSTVQVSDSEYQLPYLGSANQ 360
 QY 361 GCLPPFADVPFMIQYGYLTLNNGSOAVGRSSFCLEYPFSQMLRTGNNFTFSYTFEEVP 420
 361 GCLPPFADVPFMIQYGYLTLNNGSOAVGRSSFCLEYPFSQMLRTGNNFTFSYTFEEVP 420
 DB 361 GCLPPFADVPFMIQYGYLTLNNGSOAVGRSSFCLEYPFSQMLRTGNNFTFSYTFEEVP 420
 QY 421 FHSYASOSLDRMLNPLIDQYLYLNRTOQSGSAQNKDLLFSRGPAGMSVQPKMLP 480
 421 FHSYASOSLDRMLNPLIDQYLYLNRTOQSGSAQNKDLLFSRGPAGMSVQPKMLP 480
 DB 421 FHSYASOSLDRMLNPLIDQYLYLNRTOQSGSAQNKDLLFSRGPAGMSVQPKMLP 480
 QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMG 540
 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMG 540
 DB 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMG 540
 QY 541 MIFKESAGASNTALDVMITDEBEIKATNPVATERFGTVA VNFQSSSTD PATGDVAMG 600
 541 MIFKESAGASNTALDVMITDEBEIKATNPVATERFGTVA VNFQSSSTD PATGDVAMG 600
 DB 541 MIFKESAGASNTALDVMITDEBEIKATNPVATERFGTVA VNFQSSSTD PATGDVAMG 600
 QY 601 ALPGMWQDRDVLQGPIMAKIPHTDGHFHPSPLMGGFGLKNPPQILIKNTPVPANPPA 660
 601 ALPGMWQDRDVLQGPIMAKIPHTDGHFHPSPLMGGFGLKNPPQILIKNTPVPANPPA 660
 DB 601 ALPGMWQDRDVLQGPIMAKIPHTDGHFHPSPLMGGFGLKNPPQILIKNTPVPANPPA 660
 QY 661 EFSATKFASTITQYSTQGVSEIEMELQKENS KRMNPEVOYTSNYAKSANDFTVDNGL 720
 661 EFSATKFASTITQYSTQGVSEIEMELQKENS KRMNPEVOYTSNYAKSANDFTVDNGL 720
 DB 661 EFSATKFASTITQYSTQGVSEIEMELQKENS KRMNPEVOYTSNYAKSANDFTVDNGL 720
 QY 721 YTEPRPIGTRYLTRPL 736
 721 YTEPRPIGTRYLTRPL 736
 DB 721 YTEPRPIGTRYLTRPL 736

RESULT 4

US-10-696-261-13
 ; Sequence 13, Application US/10696261
 ; Publication No. US20040057931A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wilson, James M.
 ; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
 ; FILE REFERENCE: GNPVN.031USA
 ; CURRENT APPLICATION NUMBER: US/10/696.261
 ; PRIOR FILING DATE: 2003-10-29
 ; PRIOR APPLICATION NUMBER: US/09/807,802A
 ; PRIOR FILING DATE: 2002-02-21

PRIOR APPLICATION NUMBER: US 60/107,114
 ; PRIOR FILING DATE: 1998-11-05
 ; PRIOR APPLICATION NUMBER: PCT/US99/25694
 ; PRIOR FILING DATE: 1999-11-02
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 736
 ; TYPE: PR
 ; ORGANISM: AAV-1
 US-10-696-261-13

Query Match 100.0%; Score 3989; DB 4; Length 736;
 Best Local Similarity 100.0%; Pred. No. 9.5e-309;
 Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGYLPDMLEDMISGIRREMDLKGAPKPKANQKQODDGRGLVLPGYKYLGPFGNLD 60
 1 MAADGYLPDMLEDMISGIRREMDLKGAPKPKANQKQODDGRGLVLPGYKYLGPFGNLD 60
 DB 1 MAADGYLPDMLEDMISGIRREMDLKGAPKPKANQKQODDGRGLVLPGYKYLGPFGNLD 60
 QY 61 KGEVNAADAAALAEHDKAYDQOLKAGDNPYLRYNHADAERLQEDTSFGNLAGRAVQ 120
 61 KGEVNAADAAALAEHDKAYDQOLKAGDNPYLRYNHADAERLQEDTSFGNLAGRAVQ 120
 DB 61 KGEVNAADAAALAEHDKAYDQOLKAGDNPYLRYNHADAERLQEDTSFGNLAGRAVQ 120
 QY 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGGQPAKKRLNFGQTGSE 180
 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGGQPAKKRLNFGQTGSE 180
 DB 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTGGQPAKKRLNFGQTGSE 180
 QY 181 SVDPQOLGEPPTPAVGPPTTMASSGGA PMADNNEGADGVGNASGNMHCSTWLGDRVI 240
 181 SVDPQOLGEPPTPAVGPPTTMASSGGA PMADNNEGADGVGNASGNMHCSTWLGDRVI 240
 DB 181 SVDPQOLGEPPTPAVGPPTTMASSGGA PMADNNEGADGVGNASGNMHCSTWLGDRVI 240
 QY 241 TTSRTMALPTNNHLKQISSASTGASNDNHFGYSTPMGYPDFNRFCHFSRDMQRL 300
 241 TTSRTMALPTNNHLKQISSASTGASNDNHFGYSTPMGYPDFNRFCHFSRDMQRL 300
 DB 241 TTSRTMALPTNNHLKQISSASTGASNDNHFGYSTPMGYPDFNRFCHFSRDMQRL 300
 QY 301 INNNWGRPKRLNFKLFNIQVEYTTNDGVTIANNLSTVQVSDSEYQLPYLGSANQ 360
 301 INNNWGRPKRLNFKLFNIQVEYTTNDGVTIANNLSTVQVSDSEYQLPYLGSANQ 360
 DB 301 INNNWGRPKRLNFKLFNIQVEYTTNDGVTIANNLSTVQVSDSEYQLPYLGSANQ 360
 QY 361 GCLPPFADVPFMIQYGYLTLNNGSOAVGRSSFCLEYPFSQMLRTGNNFTFSYTFEEVP 420
 361 GCLPPFADVPFMIQYGYLTLNNGSOAVGRSSFCLEYPFSQMLRTGNNFTFSYTFEEVP 420
 DB 361 GCLPPFADVPFMIQYGYLTLNNGSOAVGRSSFCLEYPFSQMLRTGNNFTFSYTFEEVP 420
 QY 421 FHSYASOSLDRMLNPLIDQYLYLNRTOQSGSAQNKDLLFSRGPAGMSVQPKMLP 480
 421 FHSYASOSLDRMLNPLIDQYLYLNRTOQSGSAQNKDLLFSRGPAGMSVQPKMLP 480
 DB 421 FHSYASOSLDRMLNPLIDQYLYLNRTOQSGSAQNKDLLFSRGPAGMSVQPKMLP 480
 QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMG 540
 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMG 540
 DB 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMG 540
 QY 541 MIFKESAGASNTALDVMITDEBEIKATNPVATERFGTVA VNFQSSSTD PATGDVAMG 600
 541 MIFKESAGASNTALDVMITDEBEIKATNPVATERFGTVA VNFQSSSTD PATGDVAMG 600
 DB 541 MIFKESAGASNTALDVMITDEBEIKATNPVATERFGTVA VNFQSSSTD PATGDVAMG 600
 QY 601 ALPGMWQDRDVLQGPIMAKIPHTDGHFHPSPLMGGFGLKNPPQILIKNTPVPANPPA 660
 601 ALPGMWQDRDVLQGPIMAKIPHTDGHFHPSPLMGGFGLKNPPQILIKNTPVPANPPA 660
 DB 601 ALPGMWQDRDVLQGPIMAKIPHTDGHFHPSPLMGGFGLKNPPQILIKNTPVPANPPA 660
 QY 661 EFSATKFASTITQYSTQGVSEIEMELQKENS KRMNPEVOYTSNYAKSANDFTVDNGL 720
 661 EFSATKFASTITQYSTQGVSEIEMELQKENS KRMNPEVOYTSNYAKSANDFTVDNGL 720
 DB 661 EFSATKFASTITQYSTQGVSEIEMELQKENS KRMNPEVOYTSNYAKSANDFTVDNGL 720
 QY 721 YTEPRPIGTRYLTRPL 736
 721 YTEPRPIGTRYLTRPL 736
 DB 721 YTEPRPIGTRYLTRPL 736

RESULT 5

US-10-696-282-3
 ; Sequence 3, Application US/10696282
 ; Publication No. US20040057932A1

```
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPV.031USA
; CURRENT APPLICATION NUMBER: US/10/696,282
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRF
; ORGANISM: AAV-1
; US-10-696-282-3

Query Match          100.0%; Score 3989; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 9,5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MAADGTLPMWLENDLSEGIREFWMDLKPAPKPKANQOKDDGRLVPGYKYLGPFGD 60
        |||||||
DB       1  MAADGTLPMWLENDLSEGIREFWMDLKPAPKPKANQOKDDGRLVPGYKYLGPFGD 60

QY      61  KGPVNAADAALAEHDKAVDQOLKAGDNPYLRYNHADAEFOERLQEDTSFGNLSRAVFQ 120
        |||||||
DB       61  KGPVNAADAALAEHDKAVDQOLKAGDNPYLRYNHADAEFOERLQEDTSFGNLSRAVFQ 120

QY      121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNGQTGDS 180
        |||||||
DB       121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNGQTGDS 180

QY      121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNGQTGDS 180
        |||||||
DB       121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNGQTGDS 180

QY      181  SVDPDPLGEPPTPAVAGPTTMASSGGA PMADNNEGADGVNAGSNMHCDSYTLGDRV 240
        |||||||
DB       181  SVDPDPLGEPPTPAVAGPTTMASSGGA PMADNNEGADGVNAGSNMHCDSYTLGDRV 240

QY      241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300
        |||||||
DB       241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300

QY      241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300
        |||||||
DB       241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300

QY      301  INNMGFRPRKLNFLKFNIOKEVTNDGYTTIANNLSTVOVFSDEYQLPYVLGSAHQ 360
        |||||||
DB       301  INNMGFRPRKLNFLKFNIOKEVTNDGYTTIANNLSTVOVFSDEYQLPYVLGSAHQ 360

QY      301  INNMGFRPRKLNFLKFNIOKEVTNDGYTTIANNLSTVOVFSDEYQLPYVLGSAHQ 360
        |||||||
DB       301  INNMGFRPRKLNFLKFNIOKEVTNDGYTTIANNLSTVOVFSDEYQLPYVLGSAHQ 360

QY      361  GCLPPPADVFMIPOYGYTLTNNGSAVGRSSFYCLEYFSPQMLRTGNFTFSYTEEVP 420
        |||||||
DB       361  GCLPPPADVFMIPOYGYTLTNNGSAVGRSSFYCLEYFSPQMLRTGNFTFSYTEEVP 420

QY      361  GCLPPPADVFMIPOYGYTLTNNGSAVGRSSFYCLEYFSPQMLRTGNFTFSYTEEVP 420
        |||||||
DB       361  GCLPPPADVFMIPOYGYTLTNNGSAVGRSSFYCLEYFSPQMLRTGNFTFSYTEEVP 420

QY      421  FHSYVHSSOSLDBLMPPLIDQYLYLNRTQNGSASQNKDLEFSRSPAGMSVQPKXWLP 480
        |||||||
DB       421  FHSYVHSSOSLDBLMPPLIDQYLYLNRTQNGSASQNKDLEFSRSPAGMSVQPKXWLP 480

QY      421  FHSYVHSSOSLDBLMPPLIDQYLYLNRTQNGSASQNKDLEFSRSPAGMSVQPKXWLP 480
        |||||||
DB       421  FHSYVHSSOSLDBLMPPLIDQYLYLNRTQNGSASQNKDLEFSRSPAGMSVQPKXWLP 480

QY      481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHKKDEDEKFFPMSGV 540
        |||||||
DB       481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHKKDEDEKFFPMSGV 540

QY      481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHKKDEDEKFFPMSGV 540
        |||||||
DB       481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHKKDEDEKFFPMSGV 540

QY      541  MIRGKSAGASNTALDNVMTDEBEIKATNPVATERGTVAVNFQSSSTDPAATGDVHAMG 600
        |||||||
DB       541  MIRGKSAGASNTALDNVMTDEBEIKATNPVATERGTVAVNFQSSSTDPAATGDVHAMG 600

QY      541  MIRGKSAGASNTALDNVMTDEBEIKATNPVATERGTVAVNFQSSSTDPAATGDVHAMG 600
        |||||||
DB       541  MIRGKSAGASNTALDNVMTDEBEIKATNPVATERGTVAVNFQSSSTDPAATGDVHAMG 600

QY      601  ALPGMWQODRDVYLOGPIWAKIPHTDGHFHSPLMGFGLKNPPLLINXTVPANPAP 660
        |||||||
DB       601  ALPGMWQODRDVYLOGPIWAKIPHTDGHFHSPLMGFGLKNPPLLINXTVPANPAP 660

QY      601  ALPGMWQODRDVYLOGPIWAKIPHTDGHFHSPLMGFGLKNPPLLINXTVPANPAP 660
        |||||||
DB       601  ALPGMWQODRDVYLOGPIWAKIPHTDGHFHSPLMGFGLKNPPLLINXTVPANPAP 660

QY      661  EFSATKFASEITQYSTQVSEIELEWLOKENSKRMBEVOYTSNYAKSANVDFTVDNGL 720
        |||||||
DB       661  EFSATKFASEITQYSTQVSEIELEWLOKENSKRMBEVOYTSNYAKSANVDFTVDNGL 720

QY      661  EFSATKFASEITQYSTQVSEIELEWLOKENSKRMBEVOYTSNYAKSANVDFTVDNGL 720
        |||||||
DB       661  EFSATKFASEITQYSTQVSEIELEWLOKENSKRMBEVOYTSNYAKSANVDFTVDNGL 720
```

```
QY      721  YTEPRPIGTRYLTRPL 736
        |||||||
DB       721  YTEPRPIGTRYLTRPL 736

RESULT 6
US-10-696-282-13
; Sequence 13, Application US/10696282
; Publication No. US20040057932A1
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPV.031USA
; CURRENT APPLICATION NUMBER: US/10/696,282
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 736
; TYPE: PRF
; ORGANISM: AAV-1
; US-10-696-282-13

Query Match          100.0%; Score 3989; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 9,5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MAADGTLPMWLENDLSEGIREFWMDLKPAPKPKANQOKDDGRLVPGYKYLGPFGD 60
        |||||||
DB       1  MAADGTLPMWLENDLSEGIREFWMDLKPAPKPKANQOKDDGRLVPGYKYLGPFGD 60

QY      61  KGPVNAADAALAEHDKAVDQOLKAGDNPYLRYNHADAEFOERLQEDTSFGNLSRAVFQ 120
        |||||||
DB       61  KGPVNAADAALAEHDKAVDQOLKAGDNPYLRYNHADAEFOERLQEDTSFGNLSRAVFQ 120

QY      121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNGQTGDS 180
        |||||||
DB       121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNGQTGDS 180

QY      121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNGQTGDS 180
        |||||||
DB       121  AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKKRLNGQTGDS 180

QY      181  SVDPDPLGEPPTPAVAGPTTMASSGGA PMADNNEGADGVNAGSNMHCDSYTLGDRV 240
        |||||||
DB       181  SVDPDPLGEPPTPAVAGPTTMASSGGA PMADNNEGADGVNAGSNMHCDSYTLGDRV 240

QY      181  SVDPDPLGEPPTPAVAGPTTMASSGGA PMADNNEGADGVNAGSNMHCDSYTLGDRV 240
        |||||||
DB       181  SVDPDPLGEPPTPAVAGPTTMASSGGA PMADNNEGADGVNAGSNMHCDSYTLGDRV 240

QY      241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300
        |||||||
DB       241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300

QY      241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300
        |||||||
DB       241  TTSRTMALPTYNNHLKYKQISSASTGASNDNHFGYSTPWGYDFNRFCHFSPRDQRL 300

QY      301  INNMGFRPRKLNFLKFNIOKEVTNDGYTTIANNLSTVOVFSDEYQLPYVLGSAHQ 360
        |||||||
DB       301  INNMGFRPRKLNFLKFNIOKEVTNDGYTTIANNLSTVOVFSDEYQLPYVLGSAHQ 360

QY      301  INNMGFRPRKLNFLKFNIOKEVTNDGYTTIANNLSTVOVFSDEYQLPYVLGSAHQ 360
        |||||||
DB       301  INNMGFRPRKLNFLKFNIOKEVTNDGYTTIANNLSTVOVFSDEYQLPYVLGSAHQ 360

QY      361  GCLPPPADVFMIPOYGYTLTNNGSAVGRSSFYCLEYFSPQMLRTGNFTFSYTEEVP 420
        |||||||
DB       361  GCLPPPADVFMIPOYGYTLTNNGSAVGRSSFYCLEYFSPQMLRTGNFTFSYTEEVP 420

QY      361  GCLPPPADVFMIPOYGYTLTNNGSAVGRSSFYCLEYFSPQMLRTGNFTFSYTEEVP 420
        |||||||
DB       361  GCLPPPADVFMIPOYGYTLTNNGSAVGRSSFYCLEYFSPQMLRTGNFTFSYTEEVP 420

QY      421  FHSYVHSSOSLDBLMPPLIDQYLYLNRTQNGSASQNKDLEFSRSPAGMSVQPKXWLP 480
        |||||||
DB       421  FHSYVHSSOSLDBLMPPLIDQYLYLNRTQNGSASQNKDLEFSRSPAGMSVQPKXWLP 480

QY      421  FHSYVHSSOSLDBLMPPLIDQYLYLNRTQNGSASQNKDLEFSRSPAGMSVQPKXWLP 480
        |||||||
DB       421  FHSYVHSSOSLDBLMPPLIDQYLYLNRTQNGSASQNKDLEFSRSPAGMSVQPKXWLP 480

QY      481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHKKDEDEKFFPMSGV 540
        |||||||
DB       481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHKKDEDEKFFPMSGV 540

QY      481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHKKDEDEKFFPMSGV 540
        |||||||
DB       481  GPCYRQORVSKTIDNNNSFTWTGASKYNLNGRESIINPGTMAHKKDEDEKFFPMSGV 540

QY      541  MIRGKSAGASNTALDNVMTDEBEIKATNPVATERGTVAVNFQSSSTDPAATGDVHAMG 600
        |||||||
DB       541  MIRGKSAGASNTALDNVMTDEBEIKATNPVATERGTVAVNFQSSSTDPAATGDVHAMG 600
```

```
Db 541 MIFGESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFOSSSTDATGVAHMG 600
Qy 601 ALPGWVWQDRDVLVYLGPIWAKI PHTDGHFHPSPIMGFGLKNPPQILIKNTVPVAPNPA 660
Db 601 ALPGWVWQDRDVLVYLGPIWAKI PHTDGHFHPSPIMGFGLKNPPQILIKNTVPVAPNPA 660
Qy 661 EFSATKPFASFTQYSTGVSVIEIWELOKENS KRNMPVOYTSNVAKSANVDFTVDNNGL 720
Db 661 EFSATKPFASFTQYSTGVSVIEIWELOKENS KRNMPVOYTSNVAKSANVDFTVDNNGL 720
Qy 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736

RESULT 7
US-10-696-900-3
; Sequence 3, Application US/10696900
; Publication No. US20040057933A1
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; FILE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/10/696,900
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-10-696-900-3

Query Match 100.0%; Score 3989; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 9.5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAADGYLPDMLENDLSGIREMMDLKGA PKPKANQOKODDGRGLVPGYKYLGPFGNLD 60
Db 1 MAADGYLPDMLENDLSGIREMMDLKGA PKPKANQOKODDGRGLVPGYKYLGPFGNLD 60
Qy 61 KGEVNNADAALAEHDKAYDOOLKAGDNPYLRYNHADAERQERLQEDTSPFGNIGRAVFQ 120
Db 61 KGEVNNADAALAEHDKAYDOOLKAGDNPYLRYNHADAERQERLQEDTSPFGNIGRAVFQ 120
Qy 121 AKKRVLEPLGLVEGAKTAPGKKRPVBOQPOEPDSSSGIGTGQOPAKKRLNFGQTGDS 180
Db 121 AKKRVLEPLGLVEGAKTAPGKKRPVBOQPOEPDSSSGIGTGQOPAKKRLNFGQTGDS 180
Qy 121 AKKRVLEPLGLVEGAKTAPGKKRPVBOQPOEPDSSSGIGTGQOPAKKRLNFGQTGDS 180
Db 121 AKKRVLEPLGLVEGAKTAPGKKRPVBOQPOEPDSSSGIGTGQOPAKKRLNFGQTGDS 180
Qy 181 SVDPPOPLGEBPPATPAVVGPTTMA SGGA PMADNNEGADGVGNA SGWNHCDSTWLGRVI 240
Db 181 SVDPPOPLGEBPPATPAVVGPTTMA SGGA PMADNNEGADGVGNA SGWNHCDSTWLGRVI 240
Qy 241 TTSRTVALPTYNHLYKQISSASTGASNDNHYFGYSTPMGY PDPNRFCHFSRDMQRL 300
Db 241 TTSRTVALPTYNHLYKQISSASTGASNDNHYFGYSTPMGY PDPNRFCHFSRDMQRL 300
Qy 301 INNNGRRPRKRLNKLRFNIQYKEVTYNDGVTTIANNTLSTYQVSDSYQLPYVLGSAHQ 360
Db 301 INNNGRRPRKRLNKLRFNIQYKEVTYNDGVTTIANNTLSTYQVSDSYQLPYVLGSAHQ 360
Qy 361 GCLPPPADVEMIQOYGYLTLNNGSOAVGRSSFYCLEYFSPQOMRTGNNPFESYTFEEVP 420
Db 361 GCLPPPADVEMIQOYGYLTLNNGSOAVGRSSFYCLEYFSPQOMRTGNNPFESYTFEEVP 420
```

```
Qy 421 FHSYSVHQSQSIDRLMNPILIDQYLYYLNRTONQSSAQNKDLLFSRGPAGMSVOPKMLP 480
Db 421 FHSYSVHQSQSIDRLMNPILIDQYLYYLNRTONQSSAQNKDLLFSRGPAGMSVOPKMLP 480
Qy 481 GPCYRQORVSKTKDNNNSNFTWGA SKYNLNGRESIINPCTAMASHDODDKFPPMSGV 540
Db 481 GPCYRQORVSKTKDNNNSNFTWGA SKYNLNGRESIINPCTAMASHDODDKFPPMSGV 540
Qy 541 MIFGESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFOSSSTDATGVAHMG 600
Db 541 MIFGESAGASNTALDNVMTDEBEIKATNPVATERFCTVAVNFOSSSTDATGVAHMG 600
Qy 601 ALPGWVWQDRDVLVYLGPIWAKI PHTDGHFHPSPIMGFGLKNPPQILIKNTVPVAPNPA 660
Db 601 ALPGWVWQDRDVLVYLGPIWAKI PHTDGHFHPSPIMGFGLKNPPQILIKNTVPVAPNPA 660
Qy 661 EFSATKPFASFTQYSTGVSVIEIWELOKENS KRNMPVOYTSNVAKSANVDFTVDNNGL 720
Db 661 EFSATKPFASFTQYSTGVSVIEIWELOKENS KRNMPVOYTSNVAKSANVDFTVDNNGL 720
Qy 721 YTEPRPIGTRYLTRPL 736
Db 721 YTEPRPIGTRYLTRPL 736

RESULT 8
US-10-696-900-13
; Sequence 13, Application US/10696900
; Publication No. US20040057933A1
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; FILE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/10/696,900
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-10-696-900-13

Query Match 100.0%; Score 3989; DB 4; Length 736;
Best Local Similarity 100.0%; Pred. No. 9.5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAADGYLPDMLENDLSGIREMMDLKGA PKPKANQOKODDGRGLVPGYKYLGPFGNLD 60
Db 1 MAADGYLPDMLENDLSGIREMMDLKGA PKPKANQOKODDGRGLVPGYKYLGPFGNLD 60
Qy 61 KGEVNNADAALAEHDKAYDOOLKAGDNPYLRYNHADAERQERLQEDTSPFGNIGRAVFQ 120
Db 61 KGEVNNADAALAEHDKAYDOOLKAGDNPYLRYNHADAERQERLQEDTSPFGNIGRAVFQ 120
Qy 121 AKKRVLEPLGLVEGAKTAPGKKRPVBOQPOEPDSSSGIGTGQOPAKKRLNFGQTGDS 180
Db 121 AKKRVLEPLGLVEGAKTAPGKKRPVBOQPOEPDSSSGIGTGQOPAKKRLNFGQTGDS 180
Qy 181 SVDPPOPLGEBPPATPAVVGPTTMA SGGA PMADNNEGADGVGNA SGWNHCDSTWLGRVI 240
Db 181 SVDPPOPLGEBPPATPAVVGPTTMA SGGA PMADNNEGADGVGNA SGWNHCDSTWLGRVI 240
Qy 241 TTSRTVALPTYNHLYKQISSASTGASNDNHYFGYSTPMGY PDPNRFCHFSRDMQRL 300
```

```
Db      241  |||||
Qy      301  INNMGFRPRRLNFKLFNIQVEVTNDGVTIANNLSTVOVFSDBEYOLPYVLGSAHQ 360
Db      301  INNMGFRPRRLNFKLFNIQVEVTNDGVTIANNLSTVOVFSDBEYOLPYVLGSAHQ 360
Qy      361  GCLPPPADVFMIPQYGYLTINNQSQAVERSSFCLEFPSPQMLRTGNNFTFSYTEEVP 420
Db      361  GCLPPPADVFMIPQYGYLTINNQSQAVERSSFCLEFPSPQMLRTGNNFTFSYTEEVP 420
Qy      421  FHSVYASQSLDLNMP.LIDQYLYLNRTQNGSQAQNDL.FSRGSPAGMSVQPKMWLP 480
Db      421  FHSVYASQSLDLNMP.LIDQYLYLNRTQNGSQAQNDL.FSRGSPAGMSVQPKMWLP 480
Qy      481  GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
Db      481  GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
Qy      541  MIFGKESAGASNTALDNWMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAVDYHANG 600
Db      541  MIFGKESAGASNTALDNWMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAVDYHANG 600
Qy      601  ALPGMWQORDVYLQGPYIAKIPHTDGHFHPSPLMGFGKPNPOLLIKNTFVPANPPA 660
Db      601  ALPGMWQORDVYLQGPYIAKIPHTDGHFHPSPLMGFGKPNPOLLIKNTFVPANPPA 660
Qy      661  EFSATKFASTITQYSTGOVSVEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNDGL 720
Db      661  EFSATKFASTITQYSTGOVSVEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNDGL 720
Qy      721  YTEPRPIGTRYLTRPL 736
Db      721  YTEPRPIGTRYLTRPL 736

RESULT 9
US-10-496-799-5
; Sequence 5, Application US/10496799
; Publication No. US20050014262A1
; GENERAL INFORMATION:
; APPLICANT: The Trustees of The University of Pennsylvania
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alviria, Mauricio
; TITLE OF INVENTION: Adeno-Associated Virus (AAV) Serotype 9 Sequences, Vectors Contai
; FILE REFERENCE: UPN-02734PCT
; CURRENT APPLICATION NUMBER: US/10/496,799
; CURRENT FILING DATE: 2004-06-08
; PRIOR APPLICATION NUMBER: US 60/341,150
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/386,132
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 736
; TYPE: PRT
; ORGANISM: adeno-associated virus serotype 1
US-10-496-799-5

Query Match      100.0%; Score 3989; DB 5; Length 736;
Best Local Similarity 100.0%; Pred. No. 9,5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      121  AKKRVLEPGLVBEAGKTA.PGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFCOTGDSF 180
Db      121  AKKRVLEPGLVBEAGKTA.PGKKRPVEQSPQEPDSSSGIGKTGQOPAKKRLNFCOTGDSF 180
Qy      181  SVDDPOLGEPATPAVPTTMA.SGGAP.MADNNEGADGVNASGNWCHDSFTWLDGRLV 240
Db      181  SVDDPOLGEPATPAVPTTMA.SGGAP.MADNNEGADGVNASGNWCHDSFTWLDGRLV 240
Qy      241  TSTRTVALPTYNNHL.YKQISSASTGASNDNH.FGISTWGVDFENRFCHFSFPRMORL 300
Db      241  TSTRTVALPTYNNHL.YKQISSASTGASNDNH.FGISTWGVDFENRFCHFSFPRMORL 300
Qy      301  INNMGFRPRRLNFKLFNIQVEVTNDGVTIANNLSTVOVFSDBEYOLPYVLGSAHQ 360
Db      301  INNMGFRPRRLNFKLFNIQVEVTNDGVTIANNLSTVOVFSDBEYOLPYVLGSAHQ 360
Qy      361  GCLPPPADVFMIPQYGYLTINNQSQAVERSSFCLEFPSPQMLRTGNNFTFSYTEEVP 420
Db      361  GCLPPPADVFMIPQYGYLTINNQSQAVERSSFCLEFPSPQMLRTGNNFTFSYTEEVP 420
Qy      421  FHSVYASQSLDLNMP.LIDQYLYLNRTQNGSQAQNDL.FSRGSPAGMSVQPKMWLP 480
Db      421  FHSVYASQSLDLNMP.LIDQYLYLNRTQNGSQAQNDL.FSRGSPAGMSVQPKMWLP 480
Qy      481  GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
Db      481  GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
Qy      541  MIFGKESAGASNTALDNWMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAVDYHANG 600
Db      541  MIFGKESAGASNTALDNWMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAVDYHANG 600
Qy      601  ALPGMWQORDVYLQGPYIAKIPHTDGHFHPSPLMGFGKPNPOLLIKNTFVPANPPA 660
Db      601  ALPGMWQORDVYLQGPYIAKIPHTDGHFHPSPLMGFGKPNPOLLIKNTFVPANPPA 660
Qy      661  EFSATKFASTITQYSTGOVSVEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNDGL 720
Db      661  EFSATKFASTITQYSTGOVSVEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNDGL 720
Qy      721  YTEPRPIGTRYLTRPL 736
Db      721  YTEPRPIGTRYLTRPL 736

RESULT 10
US-10-873-632A-20
; Sequence 20, Application US/10873632A
; Publication No. US20060051333A1
; GENERAL INFORMATION:
; APPLICANT: ARETMAN, Alejandra E
; APPLICANT: COLOSI, Peter
; APPLICANT: LOCHRIS, Michael A.
; APPLICANT: SUROSKY, Richard T.
; TITLE OF INVENTION: AAV VIRIONS WITH DECREASED IMMUNOREACTIVITY AND USES THEREFOR
; FILE REFERENCE: 0800-0032
; CURRENT APPLICATION NUMBER: US/10/873,632A
; CURRENT FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: 60/480,395
; PRIOR FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: 60/567,310
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: 60/576,501
; PRIOR FILING DATE: 2004-06-03
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 736
; TYPE: PRT
; ORGANISM: adeno-associated virus 1
US-10-873-632A-20
```



```
Query Match      100.0%; Score 3989; DB 5; Length 736;
Best Local Similarity 100.0%; Pred. No. 9.5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MAADGYLPDMLEDNLSGIRREWMDLKPGAPKPKANQOKODDGRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDMLEDNLSGIRREWMDLKPGAPKPKANQOKODDGRGLVLPGYKYLGPFGNGLD 60
Oy 61 KGEVYNADAAALAEHDKAYDQOLKAGDNPLYRYNHADAPEORLOEDTSFGGNLGRAVFO 120
Db 61 KGEVYNADAAALAEHDKAYDQOLKAGDNPLYRYNHADAPEORLOEDTSFGGNLGRAVFO 120
Oy 121 AKKAVLEPLGLVEBGAKTAPGKKRPVEOSPOEPDSSSGIGTKGOOPAKKRLNFGQTGSE 180
Db 121 AKKAVLEPLGLVEBGAKTAPGKKRPVEOSPOEPDSSSGIGTKGOOPAKKRLNFGQTGSE 180
Oy 181 SVDPPOPUGEPBPATPAAGPTTMASSGGAPMADNNEGADGVGNASGNMHCDSTWLGDRVI 240
Db 181 SVDPPOPUGEPBPATPAAGPTTMASSGGAPMADNNEGADGVGNASGNMHCDSTWLGDRVI 240
Oy 241 TTSRTMALPTNNHLVKQISSASTGASNDNHFGYSTPMGYFDPNRFCHFSPRDMOQL 300
Db 241 TTSRTMALPTNNHLVKQISSASTGASNDNHFGYSTPMGYFDPNRFCHFSPRDMOQL 300
Oy 301 INNNWGRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTYQVPSDSSEYOLPYVLGSAHQ 360
Db 301 INNNWGRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTYQVPSDSSEYOLPYVLGSAHQ 360
Oy 361 GCLPPPADVPMIOYGYLTNNNGSOAVGRSSFCLEFPBSOMLRTGNNFTFSYFEEVP 420
Db 361 GCLPPPADVPMIOYGYLTNNNGSOAVGRSSFCLEFPBSOMLRTGNNFTFSYFEEVP 420
Oy 421 FHSSYASQSIDRLMNPILIDQYLYLNRTONQSSAQNKDILFSGSPAGMSVOPKXMLP 480
Db 421 FHSSYASQSIDRLMNPILIDQYLYLNRTONQSSAQNKDILFSGSPAGMSVOPKXMLP 480
Oy 481 GPCYRQORVSKTKTDNNNSNFTWGASKYKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Db 481 GPCYRQORVSKTKTDNNNSNFTWGASKYKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Oy 541 MIFKESAGASNTALDNVMTDEEBEIKATNPVATERFCTVAVNFQSSSTDPAATGVDHAMG 600
Db 541 MIFKESAGASNTALDNVMTDEEBEIKATNPVATERFCTVAVNFQSSSTDPAATGVDHAMG 600
Oy 601 ALPGMWQDRDVLVYOGPIWAKIPHTDGHFHSPLMGFGLKNPPQILIKTTPVPAKP 660
Db 601 ALPGMWQDRDVLVYOGPIWAKIPHTDGHFHSPLMGFGLKNPPQILIKTTPVPAKP 660
Oy 661 EFSATKFPASFTYQSTQGVSEIEMELQKENSKEWNPVEQYTSNVAKSANDFTVDNGL 720
Db 661 EFSATKFPASFTYQSTQGVSEIEMELQKENSKEWNPVEQYTSNVAKSANDFTVDNGL 720
Oy 721 YTEBRPIGTRYLRPL 736
Db 721 YTEBRPIGTRYLRPL 736

RESULT 11
US-11-145-035-20
; Sequence 20, Application US/11145035
; Publication No. US20050287122A1
; GENERAL INFORMATION:
; APPLICANT: Bartlett et al.
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/41335
; CURRENT APPLICATION NUMBER: US/11/145.035
; PRIOR FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 10/038,972
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.3
```

```
; SEQ ID NO 20
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Adeno-associated virus 1
US-11-145-035-20

Query Match      100.0%; Score 3989; DB 6; Length 736;
Best Local Similarity 100.0%; Pred. No. 9.5e-309;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MAADGYLPDMLEDNLSGIRREWMDLKPGAPKPKANQOKODDGRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDMLEDNLSGIRREWMDLKPGAPKPKANQOKODDGRGLVLPGYKYLGPFGNGLD 60
Oy 61 KGEVYNADAAALAEHDKAYDQOLKAGDNPLYRYNHADAPEORLOEDTSFGGNLGRAVFO 120
Db 61 KGEVYNADAAALAEHDKAYDQOLKAGDNPLYRYNHADAPEORLOEDTSFGGNLGRAVFO 120
Oy 121 AKKAVLEPLGLVEBGAKTAPGKKRPVEOSPOEPDSSSGIGTKGOOPAKKRLNFGQTGSE 180
Db 121 AKKAVLEPLGLVEBGAKTAPGKKRPVEOSPOEPDSSSGIGTKGOOPAKKRLNFGQTGSE 180
Oy 181 SVDPPOPUGEPBPATPAAGPTTMASSGGAPMADNNEGADGVGNASGNMHCDSTWLGDRVI 240
Db 181 SVDPPOPUGEPBPATPAAGPTTMASSGGAPMADNNEGADGVGNASGNMHCDSTWLGDRVI 240
Oy 241 TTSRTMALPTNNHLVKQISSASTGASNDNHFGYSTPMGYFDPNRFCHFSPRDMOQL 300
Db 241 TTSRTMALPTNNHLVKQISSASTGASNDNHFGYSTPMGYFDPNRFCHFSPRDMOQL 300
Oy 301 INNNWGRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTYQVPSDSSEYOLPYVLGSAHQ 360
Db 301 INNNWGRPKRLNFKLFNIQVKEVTNDGVTTIANNLSTYQVPSDSSEYOLPYVLGSAHQ 360
Oy 361 GCLPPPADVPMIOYGYLTNNNGSOAVGRSSFCLEFPBSOMLRTGNNFTFSYFEEVP 420
Db 361 GCLPPPADVPMIOYGYLTNNNGSOAVGRSSFCLEFPBSOMLRTGNNFTFSYFEEVP 420
Oy 421 FHSSYASQSIDRLMNPILIDQYLYLNRTONQSSAQNKDILFSGSPAGMSVOPKXMLP 480
Db 421 FHSSYASQSIDRLMNPILIDQYLYLNRTONQSSAQNKDILFSGSPAGMSVOPKXMLP 480
Oy 481 GPCYRQORVSKTKTDNNNSNFTWGASKYKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Db 481 GPCYRQORVSKTKTDNNNSNFTWGASKYKYNLNGRESIINPGTAMASHKODEDKFFPMGCV 540
Oy 541 MIFKESAGASNTALDNVMTDEEBEIKATNPVATERFCTVAVNFQSSSTDPAATGVDHAMG 600
Db 541 MIFKESAGASNTALDNVMTDEEBEIKATNPVATERFCTVAVNFQSSSTDPAATGVDHAMG 600
Oy 601 ALPGMWQDRDVLVYOGPIWAKIPHTDGHFHSPLMGFGLKNPPQILIKTTPVPAKP 660
Db 601 ALPGMWQDRDVLVYOGPIWAKIPHTDGHFHSPLMGFGLKNPPQILIKTTPVPAKP 660
Oy 661 EFSATKFPASFTYQSTQGVSEIEMELQKENSKEWNPVEQYTSNVAKSANDFTVDNGL 720
Db 661 EFSATKFPASFTYQSTQGVSEIEMELQKENSKEWNPVEQYTSNVAKSANDFTVDNGL 720
Oy 721 YTEBRPIGTRYLRPL 736
Db 721 YTEBRPIGTRYLRPL 736

RESULT 12
US-10-291-583-65
; Sequence 65, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvirra, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
; Sequences and Isolating No. US20030138772A1et Sequences Identif
```

```
/ FILE REFERENCE: UPN-02735USA
/ CURRENT APPLICATION NUMBER: US/10/291,583
/ CURRENT FILING DATE: 2002-11-12
/ PRIOR APPLICATION NUMBER: US 60/350,607
/ PRIOR FILING DATE: 2001-11-13
/ PRIOR APPLICATION NUMBER: US 60/341,117
/ PRIOR FILING DATE: 2001-12-17
/ PRIOR APPLICATION NUMBER: US 60/377,066
/ PRIOR FILING DATE: 2002-05-01
/ PRIOR APPLICATION NUMBER: US 60/386,675
/ PRIOR FILING DATE: 2002-06-05
/ NUMBER OF SEQ ID NOS: 120
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 65
/ LENGTH: 736
/ TYPE: PRT
/ ORGANISM: capsid protein of AAV serotype, clone AAVEVP1
US-10-291-583-65
```

```
Query Match          99.3%; Score 3963; DB 4; Length 736;
Best Local Similarity 99.2%; Pred. No. 1.le-306;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 MAADGTLPMLEBNDLSEGIREWMDLKPAPKPKANQOKODDGRGLVPGYKYLGPFGD 60
DB 1 MAADGTLPMLEBNDLSEGIREWMDLKPAPKPKANQOKODDGRGLVPGYKYLGPFGD 60
QY 61 KGEPVNAADAALAEHDKAYDQOLKAGDNPLYLRNHADAFOERLOEDTSFGNLGRAVQ 120
DB 61 KGEPVNAADAALAEHDKAYDQOLKAGDNPLYLRNHADAFOERLOEDTSFGNLGRAVQ 120
QY 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQOQPAKKRLNGQTGDS 180
DB 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQOQPAKKRLNGQTGDS 180
QY 181 SVDDPOLGEPPTPAVGPFTTASGGAPMADNNEGADGVNAGSNWCHDSITWLGDRV 240
DB 181 SVDDPOLGEPPTPAVGPFTTASGGAPMADNNEGADGVNAGSNWCHDSITWLGDRV 240
QY 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHFGYSTPWFDPFNRFHCHFSPRDQRL 300
DB 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHFGYSTPWFDPFNRFHCHFSPRDQRL 300
QY 301 INNNMGFRPKLNFKLFNIOVKEVTNDGVTIANNLITSTVOVFSOSEYOLPYVLGSAHQ 360
DB 301 INNNMGFRPKLNFKLFNIOVKEVTNDGVTIANNLITSTVOVFSOSEYOLPYVLGSAHQ 360
QY 361 GCLPPPADVFMIPQGYLTLLNNGSOAVGRSSFYCLEYFSPQMLRTGNNTFSTFEV 420
DB 361 GCLPPPADVFMIPQGYLTLLNNGSOAVGRSSFYCLEYFSPQMLRTGNNTFSTFEV 420
QY 421 FHSSYAHOSGLDLNMLPLDQYLYLNRTONOGSAGNOKDLFSRGSFAGMSVQPKWLP 480
DB 421 FHSSYAHOSGLDLNMLPLDQYLYLNRTONOGSAGNOKDLFSRGSFAGMSVQPKWLP 480
QY 481 GPCYRQORVSKTIDNNNSNFTWTGASKYLNLRRESIINPGTAMASHKDEDEKFFPM 540
DB 481 GPCYRQORVSKTIDNNNSNFTWTGASKYLNLRRESIINPGTAMASHKDEDEKFFPM 540
QY 541 MIFGKSAGASNTALDNVMTTDEBEIKATNPVATERFGTAAVNPOSSSTDPAIGDVHMG 600
DB 541 MIFGKSAGASNTALDNVMTTDEBEIKATNPVATERFGTAAVNPOSSSTDPAIGDVHMG 600
QY 601 ALPGMWQDRDYLLOGPIMAKIPHTDGHFSPMLMGFGKHPPOQLIKNTIPVAPNPA 660
DB 601 ALPGMWQDRDYLLOGPIMAKIPHTDGHFSPMLMGFGKHPPOQLIKNTIPVAPNPA 660
QY 661 EFSATKPAFTTQYSTQVSEIEMELQKENSKRANBEVOYTSYASAVDPFVNDGL 720
DB 661 EFSATKPAFTTQYSTQVSEIEMELQKENSKRANBEVOYTSYASAVDPFVNDGL 720
QY 721 YTEPRPIGTRYLRLPL 736
DB 721 YTEPRPIGTRYLRLPL 736
```

```
DB 721 YTEPRPIGTRYLRLPL 736
RESULT 13
US-10-873-632A-19
/ Sequence 19, Application US/10873632A
/ Publication No. US2006005133A1
/ GENERAL INFORMATION:
/ APPLICANT: ARBETMAN, Alejandra E
/ APPLICANT: COLOSI, Peter
/ APPLICANT: LOCHRIE, Michael A.
/ APPLICANT: SUROSKY, Richard T.
/ TITLE OF INVENTION: AAV VIRIONS WITH DECREASED IMMUNOREACTIVITY AND USES THEREFOR
/ FILE REFERENCE: 0800-0032
/ CURRENT APPLICATION NUMBER: US/10/873,632A
/ CURRENT FILING DATE: 2004-06-21
/ PRIOR APPLICATION NUMBER: 60/480,395
/ PRIOR FILING DATE: 2003-06-19
/ PRIOR APPLICATION NUMBER: 60/567,310
/ PRIOR FILING DATE: 2004-04-30
/ PRIOR APPLICATION NUMBER: 60/576,501
/ PRIOR FILING DATE: 2004-06-03
/ NUMBER OF SEQ ID NOS: 28
/ SOFTWARE: Patent in version 3.3
/ SEQ ID NO 19
/ LENGTH: 736
/ TYPE: PRT
/ ORGANISM: adeno-associated virus 6
US-10-873-632A-19
```

```
Query Match          99.3%; Score 3963; DB 5; Length 736;
Best Local Similarity 99.2%; Pred. No. 1.le-306;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 MAADGTLPMLEBNDLSEGIREWMDLKPAPKPKANQOKODDGRGLVPGYKYLGPFGD 60
DB 1 MAADGTLPMLEBNDLSEGIREWMDLKPAPKPKANQOKODDGRGLVPGYKYLGPFGD 60
QY 61 KGEPVNAADAALAEHDKAYDQOLKAGDNPLYLRNHADAFOERLOEDTSFGNLGRAVQ 120
DB 61 KGEPVNAADAALAEHDKAYDQOLKAGDNPLYLRNHADAFOERLOEDTSFGNLGRAVQ 120
QY 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQOQPAKKRLNGQTGDS 180
DB 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTQOQPAKKRLNGQTGDS 180
QY 181 SVDDPOLGEPPTPAVGPFTTASGGAPMADNNEGADGVNAGSNWCHDSITWLGDRV 240
DB 181 SVDDPOLGEPPTPAVGPFTTASGGAPMADNNEGADGVNAGSNWCHDSITWLGDRV 240
QY 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHFGYSTPWFDPFNRFHCHFSPRDQRL 300
DB 241 TTSTRTWALPTYNNHLYKQISSASTGASNDNHFGYSTPWFDPFNRFHCHFSPRDQRL 300
QY 301 INNNMGFRPKLNFKLFNIOVKEVTNDGVTIANNLITSTVOVFSOSEYOLPYVLGSAHQ 360
DB 301 INNNMGFRPKLNFKLFNIOVKEVTNDGVTIANNLITSTVOVFSOSEYOLPYVLGSAHQ 360
QY 361 GCLPPPADVFMIPQGYLTLLNNGSOAVGRSSFYCLEYFSPQMLRTGNNTFSTFEV 420
DB 361 GCLPPPADVFMIPQGYLTLLNNGSOAVGRSSFYCLEYFSPQMLRTGNNTFSTFEV 420
QY 421 FHSSYAHOSGLDLNMLPLDQYLYLNRTONOGSAGNOKDLFSRGSFAGMSVQPKWLP 480
DB 421 FHSSYAHOSGLDLNMLPLDQYLYLNRTONOGSAGNOKDLFSRGSFAGMSVQPKWLP 480
QY 481 GPCYRQORVSKTIDNNNSNFTWTGASKYLNLRRESIINPGTAMASHKDEDEKFFPM 540
DB 481 GPCYRQORVSKTIDNNNSNFTWTGASKYLNLRRESIINPGTAMASHKDEDEKFFPM 540
QY 541 MIFGKSAGASNTALDNVMTTDEBEIKATNPVATERFGTAAVNPOSSSTDPAIGDVHMG 600
DB 541 MIFGKSAGASNTALDNVMTTDEBEIKATNPVATERFGTAAVNPOSSSTDPAIGDVHMG 600
```

Qy 601 ALPGWQDRDVLVYLGPIWAKI PHTDGHFSPMLMGFGGLKPNPPQILLIKOTPVVPANPA 660
Db 601 ALPGWQDRDVLVYLGPIWAKI PHTDGHFSPMLMGFGGLKPNPPQILLIKOTPVVPANPA 660
Qy 661 EFSATKFPASFTIYQSTGVSVIEIEMELOKENS KRNPNPEVQYTSNYSKASANDFTVDNNGL 720
Db 661 EFSATKFPASFTIYQSTGVSVIEIEMELOKENS KRNPNPEVQYTSNYSKASANDFTVDNNGL 720
Qy 721 YTEPRPIGTRLYLTRPL 736
Db 721 YTEPRPIGTRLYLTRPL 736

RESULT 14
US-11-145-035-26
; Sequence 26, Application US/11145035
; Publication No. US20050287122A1
; GENERAL INFORMATION:
; APPLICANT: Bartlett et al.
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 2835/41335
; CURRENT APPLICATION NUMBER: US/11/145,035
; PRIOR FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 10/038,972
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26
; LENGTH: 736
; TYPE: PR
; ORGANISM: Adeno-associated virus 6
US-11-145-035-26

Query Match 99.3%; Score 3963; DB 6; Length 736;
Best Local Similarity 99.2%; Pred. No. 1.1e-306;
Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAADGYLPDMLEDNLSGIREMMDLKPGAPKPKANQOQODRGVLVPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDMLEDNLSGIREMMDLKPGAPKPKANQOQODRGVLVPGYKYLGPFGNGLD 60
Qy 61 KGEVFNADAAALBHDKAYDQOLKAGDNPYLRYNHADAEPQERLOEDTSPGSLGRAVFO 120
Db 61 KGEVFNADAAALBHDKAYDQOLKAGDNPYLRYNHADAEPQERLOEDTSPGSLGRAVFO 120
Qy 121 AKKVLPEPLGIVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKRLNFGQTGDS 180
Db 121 AKKVLPEPLGIVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKRLNFGQTGDS 180
Qy 181 SVPPQPLGEPAPPAVAGPTTMASSGGAPMADNNEGADGVGNASGNWCHCSTYLGDVVI 240
Db 181 SVPPQPLGEPAPPAVAGPTTMASSGGAPMADNNEGADGVGNASGNWCHCSTYLGDVVI 240
Qy 241 TTSRTIALPTYNHLYKQISSASTGASNDNHFGYSTPMGYFPDNFCHFSPRDWORL 300
Db 241 TTSRTIALPTYNHLYKQISSASTGASNDNHFGYSTPMGYFPDNFCHFSPRDWORL 300
Qy 301 INNMGRRPKRLNFKLFNIQVEKVTNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQ 360
Db 301 INNMGRRPKRLNFKLFNIQVEKVTNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQ 360
Qy 361 GCLPPPADVFMIOYGYLTLLNNGSOAVGRSSFCLEYPSPOMLRTGNFTFSYTFEEVVP 420
Db 361 GCLPPPADVFMIOYGYLTLLNNGSOAVGRSSFCLEYPSPOMLRTGNFTFSYTFEEVVP 420
Qy 421 FHSSYASQSILDRMLNPLIDQYLYLNTQOQSGAQNKDLIFSRGSPAGMSVOPKMWLP 480
Db 421 FHSSYASQSILDRMLNPLIDQYLYLNTQOQSGAQNKDLIFSRGSPAGMSVOPKMWLP 480
Qy 481 GPCYRQORVSKTKTDNNNSNFTWTGASKRLNNGRESIINPGTAMASHKDEDEKFFPMSCV 540

Db 481 GPCYRQORVSKTKTDNNNSNFTWTGASKRLNNGRESIINPGTAMASHKDEDEKFFPMSCV 540
Qy 541 MIFGESAGASNTALDNVMTDEEBIKATNPATERFGTVAVNFOSSTDPATGAVHMG 600
Db 541 MIFGESAGASNTALDNVMTDEEBIKATNPATERFGTVAVNFOSSTDPATGAVHMG 600
Qy 601 ALPGWQDRDVLVYLGPIWAKI PHTDGHFSPMLMGFGGLKPNPPQILLIKOTPVVPANPA 660
Db 601 ALPGWQDRDVLVYLGPIWAKI PHTDGHFSPMLMGFGGLKPNPPQILLIKOTPVVPANPA 660
Qy 661 EFSATKFPASFTIYQSTGVSVIEIEMELOKENS KRNPNPEVQYTSNYSKASANDFTVDNNGL 720
Db 661 EFSATKFPASFTIYQSTGVSVIEIEMELOKENS KRNPNPEVQYTSNYSKASANDFTVDNNGL 720
Qy 721 YTEPRPIGTRLYLTRPL 736
Db 721 YTEPRPIGTRLYLTRPL 736

RESULT 15
US-11-145-035-22
; Sequence 22, Application US/11145035
; Publication No. US20050287122A1
; GENERAL INFORMATION:
; APPLICANT: Bartlett et al.
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 2835/41335
; CURRENT APPLICATION NUMBER: US/11/145,035
; PRIOR FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 10/038,972
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
; LENGTH: 736
; TYPE: PR
; ORGANISM: Adeno-associated virus 3
US-11-145-035-22

Query Match 88.0%; Score 3511; DB 6; Length 736;
Best Local Similarity 86.8%; Pred. No. 1.2e-270;
Matches 640; Conservative 38; Mismatches 57; Indels 2; Gaps 2;

Qy 1 MAADGYLPDMLEDNLSGIREMMDLKPGAPKPKANQOQODRGVLVPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDMLEDNLSGIREMMDLKPGAPKPKANQOQODRGVLVPGYKYLGPFGNGLD 60
Qy 61 KGEVFNADAAALBHDKAYDQOLKAGDNPYLRYNHADAEPQERLOEDTSPGSLGRAVFO 120
Db 61 KGEVFNADAAALBHDKAYDQOLKAGDNPYLRYNHADAEPQERLOEDTSPGSLGRAVFO 120
Qy 121 AKKVLPEPLGIVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKRLNFGQTGDS 180
Db 121 AKKVLPEPLGIVEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGOQPAKRLNFGQTGDS 180
Qy 181 SVPPQPLGEPAPPAVAGPTTMASSGGAPMADNNEGADGVGNASGNWCHCSTYLGDVVI 240
Db 181 SVPPQPLGEPAPPAVAGPTTMASSGGAPMADNNEGADGVGNASGNWCHCSTYLGDVVI 240
Qy 241 TTSRTIALPTYNHLYKQISSASTGASNDNHFGYSTPMGYFPDNFCHFSPRDWORL 300
Db 241 TTSRTIALPTYNHLYKQISSASTGASNDNHFGYSTPMGYFPDNFCHFSPRDWORL 300
Qy 301 INNMGRRPKRLNFKLFNIQVEKVTNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQ 360
Db 301 INNMGRRPKRLNFKLFNIQVEKVTNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQ 360
Qy 361 GCLPPPADVFMIOYGYLTLLNNGSOAVGRSSFCLEYPSPOMLRTGNFTFSYTFEEVVP 420
Db 361 GCLPPPADVFMIOYGYLTLLNNGSOAVGRSSFCLEYPSPOMLRTGNFTFSYTFEEVVP 420
Qy 421 FHSSYASQSILDRMLNPLIDQYLYLNTQOQSGAQNKDLIFSRGSPAGMSVOPKMWLP 480
Db 421 FHSSYASQSILDRMLNPLIDQYLYLNTQOQSGAQNKDLIFSRGSPAGMSVOPKMWLP 480
Qy 481 GPCYRQORVSKTKTDNNNSNFTWTGASKRLNNGRESIINPGTAMASHKDEDEKFFPMSCV 540

Qy	42	FHSSVASHOSLBIKAMPFLIDQLYLYLRQO-NQSSAONKOLLSRSGPAMSVQPNKL	479
Db	420	FHSSVASHOSLBIKAMPFLIDQLYLYLRQOQTSTTQOSKULLSQAGPQSMQIQAANWL	479
Qy	480	PGECTYQORVSKTKTDNNNSNTTWGASKTYNNGEBSLIINGTAMASHKDEKFFPMSC	539
Db	480	PGCTYQORISKTDANNNSNFPMPWPAASRYHLNGDSLNVNGPMAASHKDEEFPFMHG	539
Qy	540	VMI-FGKESAGASNTALDNVMTDEEBEIKATNPVATEREGTAVNFOSSSTDPAQDVHAM	599
Db	540	NLI-FGKEGTTASNAELDNVMTDEEBEIRTTNPVATEQGTAVANNQSSNTAPTRTVNDQ	599
Qy	600	GALPGMWQDRVYLOGPFWAKIPTHDDGHFHSPLMGFGKLNPPQLLINTVTPANPP	659
Db	600	GALPGMWQDRVYLOGPFWAKIPTHDDGHFHSPLMGFGKLNHPQQLMINTVTPANPP	659
Qy	660	AEEFSATKFASTFYQSTGVSEIEMELQKENSKEWNEVOYTSNYAKSANVDETVDNG	719
Db	660	TTESPAKPASFTFYQSTGVSEIEMELQKENSKEWNEIYQTSYVKNKSVANVDETVDNG	719
Qy	720	LYTEPRPIGTRYLTRPL	736
Db	720	YISEPRPIGTRYLTRNL	736

Search completed: June 12, 2006, 12:57:21
Job time : 185 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2006, 12:47:29 ; Search time 45 Seconds
(without alignments)
1573.679 Million cell updates/sec

Title: US-10-696-261-13

Perfect score: 3989
Sequence: 1 MADGGYLPDWLEDNLESGIR.....NNGLYTEPRPIGRYTLRPL 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR 80:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2177.5	54.6	504	1 VCPV3A	coat protein - ade
2	2168.5	54.4	732	2 S52210	coat protein Vp1 -
3	798.5	20.0	673	1 VCPV85	coat protein Vp1 -
4	600.5	15.1	781	1 VCPV19	coat protein Vp1 -
5	482	12.1	729	1 A60006	coat protein Vp1 -
6	482	12.1	729	1 VCPVNA	coat protein Vp1 -
7	444.5	11.1	727	1 VCPV1F	coat protein Vp1 -
8	439.5	11.0	722	1 VCPVME	coat protein Vp1 -
9	431.5	10.8	727	1 VCPVFP	coat protein Vp1 -
10	426	10.7	723	1 VCPVPP	coat protein Vp1 -
11	424.5	10.6	718	1 VCPVIM	coat protein Vp1 -
12	421.5	10.6	748	1 VCPVCP	coat protein Vp1 -
13	407.5	10.2	737	1 VCPVCD	coat protein Vp1 -
14	404	10.1	722	1 VCPVVC	coat protein Vp1 -
15	385.5	9.7	722	1 VCPVCN	coat protein Vp1 -
16	377	9.5	716	1 VCPV2M	coat protein Vp1 -
17	234	5.9	587	1 B44276	coat protein Vp1 -
18	199.5	5.0	810	2 A44054	orf1 protein - Jun
19	198	5.0	584	2 S49594	capsid protein VP2
20	191	4.8	702	1 VCPVAP	coat protein Vp1 -
21	143.5	3.6	648	2 S50856	whn protein - rat
22	139	3.5	931	2 T49710	related to glucan
23	131.5	3.3	1446	1 A45344	immediate-early pr
24	130.5	3.3	1145	2 T18235	transcription acti
25	130	3.3	1338	2 T30565	MAP kinase kinase
26	129	3.2	1072	2 A86827	hypothetical prote
27	124.5	3.1	2493	2 A55481	adenylate cyclase
28	124	3.1	1296	2 C82521	hemolysin-type cal
29	124	3.1	3078	2 T28432	variant-specific s

30	123.5	3.1	690	2 AB0124	probable TonB-depe
31	123.5	3.1	2529	2 B64635	toxlin-like outer m
32	123	3.1	1142	2 T37455	enamelin precursor
33	122.5	3.1	880	1 SYBSVS	valine-tRNA ligase
34	121	3.0	635	2 F96660	protein F2K1.10 (
35	121	3.0	1113	2 S28925	nuclear pore compl
36	121	3.0	1549	2 T48103	type VII collagen
37	120	3.0	1379	2 T37752	hypothetical protein
38	119.5	3.0	1180	2 E86719	hypothetical prote
39	119.5	3.0	1394	2 S60762	IGA-specific serin
40	119.5	3.0	2004	2 AC0314	probable membrane
41	119	3.0	667	2 A41311	transcription fact
42	118.5	3.0	2399	2 H71879	toxlin-like outer m
43	118	3.0	739	2 T52289	probable transkeo
44	118	3.0	956	2 T08144	myrosinase-binding
45	117.5	2.9	1076	1 A35622	nuclear pore prote

ALIGNMENTS

RESULT 1					
VCPV3A					
coat protein - adeno-associated virus type 2					
C:Species: adeno-associated virus type 2					
C:Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004					
C:Accession: A03698					
R:Strivastava, A.; Lueby, E.W.; Berns, K.I.					
J. Virol. 45, 555-564, 1983					
A:Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.					
A:Reference number: A03694; MUID:83164299; PMID:6300419					
A:Accession: A03698					
A:Status: translation not shown					
A:Molecule type: DNA					
A:Residues: 1-504 <SRI>					
A:Cross-references: UNIPROT:P03135; UNIPARC:UPI0000127C8E; EMBL:J01901; NID:g209616; PI					
C:Superfamily: adeno-associated virus coat protein					
C:Keywords: coat protein					
Query Match	54.6%	Score 2177.5;	DB 1;	Length 504;	
Best Local Similarity	82.5%	Pred. No. 3.9e-134;			
Matches 392;	Conservative 33;	Mismatches 47;	Indels 3;	Gaps 2;	
QY	203	MASGGAPMADNNEGADGAGNAGNMGHCDSTWLGDRVITTSRTWALPTNNHLYKQISS	262		
DB	1	MATSGAPMADNNEGADGAGNAGNMGHCDSTWMDRVITTSRTWALPTNNHLYKQISS	60		
QY	263	ASTGASNDNHYFGYSTWGYDFENRPHCHSPRDMORLNNNMGRPRPKLNFIOVK	322		
DB	61	QS-GASNDNHYFGYSTWGYDFENRPHCHSPRDMORLNNNMGRPRPKLNFIOVK	119		
QY	323	EVTNDGVTIANLITSTVQVFSDEYOLPYVLGSAHQCLPPPADVFMIPQGYLTIN	382		
DB	120	EVTNDGVTIANLITSTVQVFTDSYOLPYVLGSAHQCLPPPADVFMIPQGYLTIN	179		
QY	383	NGSOAVGRSSPYCLEYRPSQMLRTGNFTFSYFEEVPHSSVYHSGSLDRLNPLIDQY	442		
DB	180	NGSOAVGRSSPYCLEYRPSQMLRTGNFTFSYFEDVPHSSVYHSGSLDRLNPLIDQY	239		
QY	443	LYYNRTQONSGSNQNDLFSRSGPAGMSQPKWMLPGPCYRQORSKTQDNNNSNFT	502		
DB	240	LYYNRTQONSGSNQNDLFSRSGPAGMSQPKWMLPGPCYRQORSKTQDNNNSNFT	299		
QY	503	WTGASKYNLNGRESIINPGTAMASHKDDKDFPMSGVMFGKESAGASNTALDNVMTD	562		
DB	300	WTGATKYNLNGRDLNVP--AMASHKDDKDFPMSGVLIRGKQSEKTVNIEKVMITD	357		
QY	563	EEETKATNPVATEFPGYVAVNFGSSSTDPAIGDVHAGALFGMWODRDVYLOGPYAKI	622		
DB	358	EEETGTPNPAVTEYGVSTVLQGRNQAADVNTGQVLEGMWODRDVYLOGPYAKI	417		
QY	623	PHTDGHFSPFLMGFGGLKQNPPOILIKNTPVNPANPAEFATKFAFPIITOYSTG	677		

Db 418 PHTDGHFHPBLMGFGGLKHPPOILLIKNTVPANBSTTSAKFAFITQYSTG 472

RESULT 2
S52210
coat protein VP1 - muscovy duck parvovirus
N/Alternate names: VP1 protein
C/Species: muscovy duck parvovirus
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C/Accession: S52210
R/Zadori, Z.; Erdel, J.; Nagy, J.; Kisari, J.
submitted to the EMBL Data Library, September 1993
A/Reference number: S52209
A/Accession: S52210
A/Molecule type: DNA
A/Residues: 1-732 <2AD>
A/Cross-references: UNIPROT:Q68289; UNIPARC:UP100000Fc5D; EMBL:X75093; NID:9609091; PII
A/Experimental source: strain FM
C/Genetics:
A/Gene: VP1
C/Superfamily: parvovirus coat protein
C/Keywords: coat protein

Query Match 54.4%; Score 2168.5; DB 2; Length 732;
Best Local Similarity 55.2%; Pred. No. 2.6e-133;
Matches 418; Conservative 90; Mismatches 186; Indels 63; Gaps 13;

9 DWLEDNLSGEGIRWMDIKPGAPKPKANQOQO-----DGRGLVLPYKYYLGPFN 57
10 DWYE-----TAAASRHLKAGAPKPKSNQOSQSVSTDRKPKRKNNRGFVLPYKYVGPFN 65
58 GLDKGEVNAADAALAEHDKAYDOOLKAGDNPYLRYNHADAERLOEDTSTGNGRA 117
66 GLDKGPVNAADVALEHDKAYDOOLKAGDNPYIKFGHADOEFIDNLQGTSTFGNGSKA 125
118 VFOAKRVLLEPLGLIVEGAKTAPGKRRPVQSPQEPSSSGIGTGOQPAKKRLNFGQTG 177
126 VFOAKRRIIEPLGLIVEPVNTAPAKK-----SSG-KLTDDPIVYKPKJLSE- 170
178 DSESVDPDPLG-----EPPATPAVGFPTTASGGGAPMADNNEGADGVNASG 226
171 --ENSPSPNSGGEASAAATEGSEPVAA-----NMAEGSGSGAMGDSAGADGVNASG 222
227 NMHCDSWIGDRVITTSRTMALPTYNHLYKQSSASTASDNHFGYSTPVGVDEN 266
223 NMHCDSQWLDGTYITKTRTWLPSYNNHMYQALTSGTSPDSN-TQYAGYSTPWGFDEN 281
287 RFCHGFSPDMORLNNMGFRPRKLNFKLNIQVKEVTNDGVTIANNLSTVQVFS 346
282 RFCHGSPDMORLNNHMGIRPKALFKLINVQVKEVTNODQKTIANNLSTIQVFTD 341
347 SEYQLPVIGSAHQGLPRFPADVFMIPQGYLTLLN--NGSQAVGSSFFYCLEYPPSQM 403
342 NEHQLPVLGSAATEGTWPFPSPDVYALPQGYCTMHQNGAGFNDRSADFCELYPPSQM 401
404 LRFGNNTTEYTEEEVFFHSSVYASQSLDRLMPLIDQYLYLNKRTQNGSGAKNQLLF 463
402 LRTGNNEFSFEFEVFFHSMFAHSQDLRLMPLIDQYLYNFEV-NGGRNAD----F 455
464 SRSPSPAGMSVQPKNMLPGPCYRQORVSKTK--TDNNNSNFTWTGASKYNIINGEESINPG 521
456 KKAVKAGFGAMGRNMLPGPKLLDORVAYAGGTDNVANMSIWSGKNVFLKDRYLLQPG 515
522 TAMASHKDEDEKFFPMGVMIFGKE--SAGASNTALDNVMTDEEIKATNPVATERFGT 579
516 PVATHTTEDQASSVPAQNIIGIADPYRSSTLAGISDINWTDQOIAPTNGVGMRYGL 575
580 VAVNFQSSSTDPATGDVNHAMGALPGMWQORDVYLOGPVAKTPTHGHHPSPLMGFG 639
576 TVTNEQNTTAPTAFAELVGLALPGMWQNRDYLQGPVIAKIPKTGKHPSPNLDGFG 635
640 LKNPPOLLIKNTVPANPPAEFSATFASFITQYSTGVSEIEMELQKENSGRNPEV 699

Db 636 LHNPPQVFIKNTVPADPPLLEYVNOKMNSYITQYSTGQCTVEWMWELRKENSKRWNPFI 695

Query 700 QYTSNYAKSANDFTVDNNGLYTERPPICGRYITREL 736
Db 696 QFTSNFGNRTSTWFAFNETGGYEDRLIGRILYLONL 732

RESULT 3
VCPVB5
coat protein VP1 - bovine parvovirus
N/Contains: coat protein VP2
C/Species: bovine parvovirus
C/Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C/Accession: A26104
R/Chen, K.C.; Shull, B.C.; Moses, E.A.; Lederman, M.; Stout, E.R.; Bates, R.C.
J. Virol. 60, 1085-1097, 1986
A/Title: Complete nucleotide sequence and genome organization of bovine parvovirus.
A/Reference number: A26104; WUID:87061184; PMID:3783814
A/Accession: A26104
A/Molecule type: DNA
A/Residues: 1-673 <CHE>
A/Cross-references: UNIPROT:P07297; UNIPARC:UP10000127D77; EMBL:M14363; NID:9333454; PI
C/Superfamily: parvovirus coat protein
C/Keywords: coat protein
F/138-673/Product: coat protein VP2 #status predicted <VP2>

Query Match 20.0%; Score 798.5; DB 1; Length 673;
Best Local Similarity 28.6%; Pred. No. 3.7e-44;
Matches 214; Conservative 105; Mismatches 275; Indels 155; Gaps 21;

30 PKPRKNOQKODDERGLVLPYKYYLGPFGGLDKGEVNAADAALAEHDKAYDOOLKAGDNP 89
3 PTMKANSK-----GLTLRYNVLGFPNLSLFAAPAPNKADAARKHDFGSDLLKGGKNP 57
90 YLRYNHADAERLOEDTSTGNGRAVFOAKKRVLEPLGLIVEGAKTAPGKRRPVQSP 149
58 YLYFNTHDQNLIDELKDDTSFGGKLARGVFOIKKALPAL-----FGTSKGDRLAKRK 111
150 PQEPDSSSGIGTGOQPAKKRLNFGQTGDSVPDQPLGEP--PATPAVGFPTTASGG 207
112 LYFARSNKGAKKANREPAPSTSNQNMVSNIDIPNDAGNQLIELATRSVVGSGSGVG 171
208 GAPMADNNEGADGVNASGMHCDSTWIGDRVITTSRTMALPTYNHLYKQISSASTGA 267
172 -----RGSGVGTSTGWTGTLFSENIYVTKNTRQICIKNGHLKS-EVLTG 222
268 SUNDNHYFGYSTPWGYDFENFCHFSPPDMORLNNMGFRPRKLNFKLNIQVKEVTN 327
223 TAHROY-ALTTPMSYFNFNQYSSHFSPPDMQHLVNDYERFRPRAMIVRVYNLQIKQIMTD 281
328 DGYTTIANN-LTSTVOVFSSEYQLPVIGSAHQGLPRFPADVFMIPQGYLTLLN----- 381
282 GAAGTYVNNDLTAGMHI FCDGDRHYPVQHPWDQCPBELPNSIWELEPQYAVIPAPISV 341
382 -NNGSQAVGSS-----SFYCLEYFPSPQMLRTGNNFTSYTFEEVFFHSSVYASQSLDRLM 435
342 DNNITTVTEHLLKGVPLVYMLENSDHEVLNNG-----RIT 376
436 NPLIDQYLYLNFTQN-----QSGSAQNQLLFSRGP-----AGMSVQPKWLP 480
377 RIVYQLWRLMDRKQHHI QHASDVOSTGQKQKLLIQRTKQPNKORFQNAALRTSWMMS 436
481 GPCTYRQORVSKTITDNNNSNFTWTGASKYNIINGEESINPGTAMASHKDEDEKFFPMG 540
437 GP-----GLARGTHNATLTQSGAGALVTWYTNQAD--VSGV 470
541 M-----IFGKE-----SAGASNTALDNVMTDEEIKATNPVATE 575
471 RAVRVGYSTDPITYGQGPESDILRLKRSASAAAGQGNPILEN-----AABH 516
576 RFQTVAVNFQSSSTDPATGDVNHAMGALPGMWQORDVYLOGPVAKTPTHGHHPSPLMGFG 635
517 TPTREATKLTITGSGADGDYKEMWMLPNQMWDSAPISRYNPIMVKVPRVVRKTLTD 576


```
Db      563 QAPLNLNTNNNTGTLPSDPLGGKRNHFMMTLNTYGPLTALNNTAPVPFNGQIWDELD 622
      626 D--GHFSPPLMGFGFLK-NPPQILIKNTVPANPPAEFSA-TKFASFITYOSTGVSV 681
Db      623 DLKPRH--VTPAFVCKNNPPGQLFVKIAP---NLTDPFNADSPQOPRIITYSNFWK 676
      682 ELEMELQKENSKRWNPEVOYTSNYSKASANDFTVDNNGLYTEPRPIG 728
      677 TLTFIAKMRSSNMNMPIQOHTT-----TAENIGNYI-PTNIG 712

RESULT 6
VCPVNA
coat protein VP1 - porcine parvovirus (strain NADL-2)
N:contains: coat protein VP2
C:Species: porcine parvovirus
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: B33743; D48472
R:Vasudevacharya, U.; Basak, S.; Srinivas, R.V.; Compans, R.W.
Virology 173, 368-377, 1989
A:Title: Nucleotide sequence analysis of the capsid genes and the right-hand terminal pa
A:Reference number: A33743; MUID:90085785; PMID:2596019
A:Accession: B33743
A:Molecule type: DNA
A:Residues: 1-729 <VAS>
A:Cross-references: UNIPROT:P18546; UNIPARC:UPI0000127C70; GB:M32787; NID:G332983; PIDN:
R:Bergeron, J.; Menezes, J.; Tjissen, P.
Virology 197, 86-98, 1993
A:Title: Genomic organization and mapping of transcription and translation products of p
A:Reference number: A48472; MUID:94025614; PMID:8212558
A:Accession: D48472
A:Molecule type: DNA
A:Residues: 11-729 <BER>
A:Cross-references: UNIPARC:UPI0000174965
A:Experimental source: NADL-2, ATCC VR-742
A:Note: Sequence extracted from NCBI backbone (NCBIN:138789, NCBI:P138794)
C:Genetics:
A:Intons: 10/1
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein; glycoprotein
F:151-729/Product: coat protein VP2 #status predicted <VP2>
F:12,172,198,282,330,433,441,573,604,651/Binding site: carbohydrate (asn) (covalent) #st

Query Match      12.1%; Score 482; DB 1; Length 729;
Best Local Similarity 25.5%; Pred. No. 1,7e-23;
Matches 197; Conservative 111; Mismatches 310; Indels 154; Gaps 35;

Qy      43 RGLVLPQYKYLGPFGNGLDKPEPVNAADAALEHDKADQOLKGDNPYLRYNHADAFOE 102
      9 RGLTLPQYKYLGPFGNSLDGEPNPSDAAKEHDEADKYIKGKPNFYFSAADKRFIK 68
      103 RLQEDTSPFGNLRGAVFQAKRVLEPLGLVEBGAKTAPGKKRPEVDSPOEPSSSGIGK- 161
      69 ETEHADYGKIKGHYFRAKARAPKL---SETDPSPTTSQOPEVRARSPKRGKSPKRGK 125
      162 -----TGQOPAKKRLNFGOTGSESVDPDPLGEPAPATPAVGPPTTMAAGGAPMA 212
      126 PABRHIFINLAKKAKGTSNTNSMSSENEOHPIINAGTEL-SATNESGGGGG--- 181
      213 DNNEGADVGNAAG--NMHCDSITWLGD---RVITTSRTMALPTYYNNHLYKQIS--SAST 265
      182 -GGRGAGGVGVSSTGFNNQTEFQYLGLVRLTAHAARLTHLNNPHEETKRIHVNLNSES 240
      266 GAS-----NDNHFGYSTPFGVFDENRFCHFSPPDMQRLINNMGFRPKSLNFKLPIQV 321
      241 GVAQGVQVQDAHQMTVPMSLIDANAGVWPNPDMQDLISNMTEINLVSEFQEIPIVVL 300
      322 KEVT---TNDGVTTIANLTSTVQVFSDSYQLPYVLGSAHQCLPPFPADEVMIPOGY 378
      301 KTTESATSPPTKIKYNNDLTASLWALDTNNLTLPYTPAAPRESETLGFYPMLPKTPQYRY 360
      379 L-----TLNNGSQAVGRS-----SFYCLE-YFPQMLATGANNF-TFSITFE 417
```

```
Db      361 YLSCIRNLNPPYTGSGQQITDSIGTGLSDIMFYILENAVPHLRLTGDEFSTGIYHP 420
      418 EVFESSVYASOSLDR-----LNNPLI--DQVLYNIRTONOGSAGNKLFSRGS 467
      421 TNPPL--KLHNSQNTNSLSGLPRPLTEPTBEGQHGTLPAAVTRGKYHQTINNSYTEAT 478
      468 PAKMSVQPKMLPGPCYRQORVSKITDNNNSFTYTGASKVYLNREGSLIN--GTAMA 525
      479 ---AIRP-----AQGVNTPYNNFEYSNGGF-----LTPIVPTADT 512
      526 SHKDDKDFPMGGMVIMFGESAGASNTALD---NMYITDEEIK--ATNP-----VAT 574
      513 QVNDDEPN-----GALRFMDYHGHLITSSQLEKTEFNPQSKGRAPK 557
      575 ERFQIVA-VNFQSS-----TDPATG--DVHMGALP-----GMVQDRDYVLQEIWA 620
      558 QQFNQQAPLNLNTNNNTGTLPSDPLGKSNHFMNLTNTYGPLTALNNTAPVPFNGQIWD 617
      621 KLPHTD--GHFSPPLMGFGFLK-NPPQILIKNTVPANPPAEFSA-TKFASFITYOST 676
      618 KELDTDLKPRH--VTPAFVCKNNPPGQLFVKIAP---NLTDPFNADSPQOPRIITYSN 671
      677 GQVSEIEMELQKENSKRWNPEVOYTSNYSKASANDFTVDNNGLYTEPRPIG 728
      672 FMMKGTILTFIAKMRSSNMNMPIQOHTT-----TAENIGNYI-PTNIG 712

RESULT 7
VCPVIF
coat protein VP1 - feline panleukopenia virus
N:contains: coat protein VP2
C:Species: feline panleukopenia virus, FPLV
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C:Accession: A03701
R:Carlson, J.; Rushlow, K.; Maxwell, I.; Maxwell, F.; Winston, S.; Hahn, W.
J. Virol. 55, 574-587, 1985
A:Title: Cloning and sequence of DNA encoding structural proteins of the autonomous par
A:Reference number: A03697; MUID:85265017; PMID:2591581
A:Accession: A03701
A:Molecule type: DNA
A:Residues: 1-727 <CAR>
A:Cross-references: UNIPROT:P04864; UNIPARC:UPI0000127D55; EMBL:M10824; NID:G333474; PI
C:Genetics:
A:Intons: 11/1
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match      11.1%; Score 444.5; DB 1; Length 727;
Best Local Similarity 23.7%; Pred. No. 4.6e-21;
Matches 187; Conservative 117; Mismatches 295; Indels 189; Gaps 35;

Qy      43 RGLVLPQYKYLGPFGNGLDKPEPVNAADAALEHDKADQOLKAGDNPYLRYNHADAFOE 102
      10 RGLVPEQYKYLGPFGNSLDGEPNPSDAAKEHDEAAALRLRGKQNPYLFSPADQRFID 69
      103 RLQEDTSPFGNLRGAVFQAKRVLEPLGLVEBGAKTAPGKKRPEVDSPOEPSSSGIGKT 162
      70 QTKDADTWGKIKGHYFRAKKAIAVLTDPDHPDTSRPTK-PTKSKPPPHIFINLAK- 127
      163 GQOPAKKRLNFGOTGSESV---DPPLGEPRA--TPAAVGPPTTMAAGGAPMADNNE 216
      128 -----KKKAGAGVVKEDNQA PMSDGA VQDPGGQPAVNERATSGNGSGGG-----GG 176
      217 GADGVNAGSNMH-----CDSITWLGDRVITTSRTMALPTYYNNHLYKQI-----SS 262
      177 GSGGAGISTGFNNQTEFRLKENGWV--ETTAISSRLVHLMPESEYKAVVYVNMMDKTA 234
      263 ASTGASNDNHYFGYSTPFGVFDENRFCHFSPPDMQRLINNMGFRPKRLNFKLPIQV 322
      235 VKGNMALDTHVQIVTFPWSLVVDANAGVWPNPDMQDLIVTMSIELHLVSEFQEIFVNLVK 294
```

```
QY 323 EVT---TNDGVTIANNTSTVOYFSDSEYOLPYVLGSAHQCL-----PPPPA----- 368
Db 295 TVSESATQPTPKVYNNDLTASLWALDSNNMTPTPAMRSETLGFYPMKPTLPTPMWRY 354
QY 369 ---DVFMIPQGYLTLNNGSOAVGRS-----SFYCLE-YFPSQMLRTGNF-TFS 413
Db 355 FQMDRLTLPFH-----TGTSGPTNTYHGTDPPDDVOFTIENSVPHLRTGDEFAITGT 408
QY 414 YTEEEVPHSSVAHSQSLDRIMN--PLIDQYLYLNRTONOGSAGKNDLLFSRSGPAGM 471
Db 409 FFFDCKP--CRLHTWQTNRALGLRP-----FLNSLPQSEGATNPGDIGVQDCKRGV 459
QY 472 S-----VQPKN-WLPGPCYRQORVSK-----TKTDNNNSNFTWTG 504
Db 460 TQMGNTDYITBATIMRPAEVGYSAPYVSFEASTQGPFTPIAAGRGAGQTDENQAA---D 516
QY 505 GASKY---NLNGRESIINPGT---AMASHKDDDEKFFPMGCVMTFGKESAGASNTALD- 556
Db 517 GDRYVAFGRHQGQKTTTGTTERFTYIAHDT-----GRYPAGDWIQNINFN 563
QY 557 NVMITDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWODRDVYLQ 616
Db 564 NLPTNDNVLLPTDPDG-----GKTGINY--TNIFNTYGPLTALNNV-----PVYPNG 610
QY 617 PIAAKLPHITDGHFHPSPLMGFGGLK-----NPPQILIKNTVPVAN--PPA 660
Db 611 QIMDKERTD-----LKPRLVHVAAPFVQNNCPGQLFVKVAENLTNEYDPA 657
QY 661 EFATKFPASFTQYSGQVSVEIEMELQKENSKRNPVEVOYTSNVAKANVDFTVNNGL 720
Db 658 SANMSR---IVTYSDFWPKGLVFKAKLRASHTNPIQWMSIN-----VDNQFN 703
QY 721 YTEPRPIG 728
Db 704 YL-PNNIG 710
```

RESULT 8

coat protein VP1 - mink enteritis virus (strain Abashiri)
N:Contains: coat protein VP2
C:Species: mink enteritis virus, MEV
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: B38350
R:Kariatsumari, T.; Horiuchi, M.; Hama, E.; Yaguchi, K.; Ishiguro, N.; Goto, H.; Shinas
J. Gen. Virol. 72, 867-875, 1991
A:Title: Construction and nucleotide sequence analysis of an infectious DNA clone of the
A:Reference number: A38350; MUID:91202123; PMID:2016597
A:Accession: B38350
A:Molecule type: DNA
A:Residues: 1-722 <NAR>
A:Cross-references: UNIPROT:P27437; UNIPARC:UPI0000174967; GB:D00765
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:139-722/Product: coat protein VP2 #status predicted <VP2>

Query Match 11.0%; Score 439.5; DB 1; Length 722;
Best Local Similarity 23.6%; Pred. No. 9.6e-21;
Matches 186; Conservative 117; Mismatches 225; Indels 189; Gaps 35;

```
QY 44 GLVLPQGYKYLPNGLDKGEPPVNAADAALHDKAVDQQLKAGDNPYLRYNHADAEPQ 103
Db 6 GLVLPQGYKYLPNGSLDQGEPTNPSDAAAKEHDEAVAAVLRSGKNPYLYFSPADQRFID 65
QY 104 LOEDTSFGNULGRAVFOAKKRVLEPLGLVEBAKTAQKRRPVQSPQEDSSSGIGKT 163
Db 66 TKQATPDGKTIHYFFPAKKAIAVLTDPDNPSTSPPTK-PTKSKRPPIHIFINLAK-- 122
QY 164 QQPAKRLNFGQTDSESEVP-----DPOPLGEPPA--TPAAVGPTTVAASGGAPWADNEG 217
Db 123 ----KKKAGAGQVYRDLAWMSDCAVQPDGGQPAVVRERATNGSNGSGGGG-----GGG 172
QY 218 ADGVGNAASGNW-----CDSTWLGDRVITTTTWTALPTYNHNLTKQI-----SSA 263
```

```
Db 173 SGGVSGISTGFNNQTEKFLENGWV--ELTANSLRLVHLNPPSENTKRVVANNMDKTAV 230
QY 264 STGASNDNHVFGYSTPMGFEDFNHCHFSBRDQRLINNNMGPRPKRLNFKLPTNIOVKE 323
Db 231 KGNMALDHTVQIVTPTMSLVDAANMGVFPBGDQLVNTMSELHIVSFBDEINNVLYKT 290
QY 324 VT---TNDGVTIANNTSTVOYFSDSEYOLPYVLGSAHQCL-----PPPPA----- 368
Db 291 TVSESATQPTPKVYNNDLTASLWALDSNNMTPTPAMRSETLGFYPMKPTLPTPMWRY 350
QY 369 ---DVFMIPQGYLTLNNGSOAVGRS-----SFYCLE-YFPSQMLRTGNF-TFS 414
Db 351 QMDRLTLPFH-----TGTSGPTNTYHGTDPPDDVOFTIENSVPHLRTGDEFAITGT 404
QY 415 YTEEEVPHSSVAHSQSLDRIMN--PLIDQYLYLNRTONOGSAGKNDLLFSRSGPAGM 472
Db 405 FFFDCKP--CRLHTWQTNRALGLRP-----FLNSLPQSEGATNPGDIGVQDCKRGV 455
QY 473 -----VQPKN-WLPGPCYRQORVSK-----TKTDNNNSNFTWTG 505
Db 456 QMGNTDYITBATIMRPAEVGYSAPYVSFEASTQGPFTPIAAGRGAGQTDENQAA---DG 512
QY 506 ASKY---NLNGRESIINPGT---AMASHKDDDEKFFPMGCVMTFGKESAGASNTALD-N 557
Db 513 DPRAVAFGRHQGQKTTTGTTERFTYIAHDT-----GRYPAGDWIQNINFN 559
QY 558 VMITDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWODRDVYLQ 617
Db 560 LPPTNDNVLLPTDPDG-----GKTGINY--TNIFNTYGPLTALNNV-----PVYPNG 606
QY 618 IMAKIPHTDGHFHPSPLMGFGGLK-----NPPQILIKNTVPVAN--PPA 661
Db 607 IMDKEFTD-----LKPRLVHVAAPFVQNNCPGQLFVKVAENLTNEYDPA 653
QY 662 FSATKFPASFTQYSGQVSVEIEMELQKENSKRNPVEVOYTSNVAKANVDFTVNNGL 721
Db 654 ANMSR---IVTYSDFWPKGLVFKAKLRASHTNPIQWMSIN-----VDNQFN 699
QY 722 TEPRPIG 728
Db 700 L-PNNIG 705
```

RESULT 9

coat protein VP1 - feline panleukopenia virus (strain 193)
N:Contains: coat protein VP2
C:Species: feline panleukopenia virus, FPLV
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C:Accession: B36608
R:Marlyn, J.C.; Davidson, B.E.; Studdert, M.J.
J. Gen. Virol. 71, 2747-2753, 1990
A:Title: Nucleotide sequence of feline panleukopenia virus: comparison with canine par
A:Reference number: A36608; MUID:91073139; PMID:2174965
A:Accession: B36608
A:Molecule type: DNA
A:Residues: 1-727 <NAR>
A:Cross-references: UNIPROT:P24840; UNIPARC:UPI0000127D56; GB:X55115
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:144-727/Product: coat protein VP2 #status predicted <VP2>

Query Match 10.8%; Score 431.5; DB 1; Length 727;
Best Local Similarity 23.5%; Pred. No. 3.2e-20;
Matches 185; Conservative 117; Mismatches 297; Indels 189; Gaps 35;

```
QY 43 RGLVLPQGYKYLPNGLDKGEPPVNAADAALHDKAVDQQLKAGDNPYLRYNHADAEPQ 102
Db 10 RGLVLPQGYKYLPNGSLDQGEPTNPSDAAAKEHDEAVAAVLRSGKNPYLYFSPADQRFID 69
QY 103 RUEDTSFGNULGRAVFOAKKRVLEPLGLVEBAKTAQKRRPVQSPQEDSSSGIGKT 162
Db 103 RUEDTSFGNULGRAVFOAKKRVLEPLGLVEBAKTAQKRRPVQSPQEDSSSGIGKT 162
```

Db 70 QTKDAKMGCKIGHYFRKAKAIAPLVLTTPDHSTSRPTK-PTKSKRPPIHFIINLAK- 127
QY 163 GOOPAKRNLNFGQTGDSSEVP-----DPOPLGEPPA--TPAAVPTTMASSGGAAPMADNNE 216
Db 128 -----KKKAGAGQVKRDNLAAPMSDGAIVQDGGQPAVRNERATGSGSGGG-----CG 176
QY 217 GADGAGNAGNNH-----CDSTWLGDRVITTTSTWMLPTYNHLYKQI-----SS 262
Db 177 GSGGVAISTGTENNQTEPFKLENGWV--ETTNASSSLVHLNMESESNYKRVVNNNDKTA 234
QY 263 ASTGASNDHNYFGYSTPWCYFDPNRFCHSPSPDMQRLINNNGWFRPKRLNFKLFNIQYK 322
Db 235 VAGNNALDIIHQIYTPMSLVDANANGWAFNPGDWLIYNTMSELHLVSEGEIPIRVVLK 294
QY 323 EYT---TNDGVTIANNLSTVQVFSDEXQLPYVLGSAHQGL-----PPPPA----- 368
Db 295 TVSESAIQPTPKYVNNDLTASLWALDNNMTPEPTAAMRSETLGFYPMKPTPIPTMRYX 354
QY 369 ---DVFMIPOYGLTLNNGSQAVGRS-----SFYCLE-YFPSSQMLRTGNNE-TFS 413
Db 355 FQWDRTLIPSH-----TGTSGTPTVYHGTDPDVQFYTTIENSVPVHLRTGDEPATGT 408
QY 414 YTFEEVPFHSSYAHQSGLRLNM--PLIDQYLYLNRTQNGSGAQNKDLFSRGSAGW 471
Db 409 FPFDCRP--CLLHTWQTRALGLP-----FLNSLPSGBATVPGLIGVQODKRGV 459
QY 472 S-----VQPKN-MLPGCYRQORVSK-----TKTDNNNSNFTWT 504
Db 460 TQMGNTDIYTEATIMRPAEVGSAFYVSFEASTQGFKPIAAGRGAGQTDENQAA---D 516
QY 505 GASKY---NLNGRESIINCT---AMASHKDEDEKFFPMGSMITGKSAGASNTALD- 556
Db 517 GQPRVAFGRHOGKTTTGTGTPRFTYIAHDT-----GRYPEGWIONINE 563
QY 557 NWMITDEEIKATNPATERFGTVAVNFQSSSTDPATGVDHAMGALPGMWQDRDYVLOG 616
Db 564 NLPVTMDNVLLPFDPLG---GKTGINY--TNLFNTYGLTALANNP-----PYTPNG 610
QY 617 PIWAKIPIHTDGHFHPSPILMGFGFLK-----NPPQILIKNTVPVAN---PPA 660
Db 611 QIWDKEFTD-----LKPRLHVNAPFVCCNCPGLFVYVAPNLITNEVDPA 657
QY 661 EFSATCFASFIITQYSTGVSEVLEIEMLOKENSKRMPDEVQYTSNYAKSANVDTVUNNG 720
Db 658 SAAMSR---ITVYSDFPMWKGLVFRKAKLASHMTNPIDQMSIN-----VDNQFN 703
QY 721 YTEPRPIG 728
Db 704 YV-PNNIG 710

RESULT 10
VCPVP
coat protein VP1 - porcine parvovirus
N:contains: coat protein VP2
C:Species: porcine parvovirus
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C/Accession: B33302
R:Ranz, A.I.; Mancius, J.J.; Diaz-Aroca, E.; Casal, J.I.
J. Gen. Virol. 70, 2541-2553, 1989
A>Title: Porcine parvovirus: DNA sequence and genome organization.
A:Reference number: A33302; MUID:90010964; PMID:2794971
A:Residues: 1-723 <RAN>
A:Molecule type: DNA
A:Cross-references: UNIPROT:P18546; UNIPARC:UP10000174964; EMBL:DD0623
C:Genetics: 10/1
A:Introns: 10/1
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
F:145-723/Product: coat protein VP2 #status predicted <VP2>

Query Match 10.7%; Score 426; DB 1; Length 723;

Best Local Similarity 24.9%; Pred. No. 7, 3e-20;
Matches 190; Conservative 109; Mismatches 308; Indels 156; Gaps 36;
QY 53 LGPFNGLDKGEVNNADAALAEHDKAYDQALAGDNPYLRYNHDAEFOERLQEDTSFGG 112
Db 13 LGPFGNSLDGEPPTNPDAAKHEDEAYDKYIKSGKNPTFYFAADEKFIKETHARDYQ 72
QY 113 NIGRAVQAKKAVLEPLGLVEGAKTAPGKRPVDSPOEDSSSIGK-----T 162
Db 73 KIGHYFFRAKRAPAPKL-----SETDSEPTTSQOEVRSPKHPQSKPFGKRAPAPRHIFINL 129
QY 163 GOOPAKRNLNFGQTGDSSEVPDPOPLGEPPA--AAVGPPTTMASSGGAAPMADNNEGADV 221
Db 130 AKKXAKGTSNTNSNSSENVQHNPIIN--AAITELSTGHSSEGGGG-----GGRAGGV 183
QY 222 GNASGNH--CDSTWIGD--RVITTSRTWMLPTYNHLYKQI-----SSASTGAS-ND 270
Db 184 GYSTGSFNNQTEFQYLGEGGLVRIITAHASRLIHLNMEHEHTYKRIHVLNSESAGQMVOD 243
QY 271 NHYFGYSTPWGYFDFNRFCHSPSPDMQRLINNNGWFRPKRLNFKLFNIQYKEVT---TN 327
Db 244 DAHTQMTPEWSLIDANAMGVWFNPADWOLISNNMTEINLVSPBOALFNVVLKTTESATS 303
QY 328 DGVTTIANNLSTVQVFSDEXQLPYVLGSAHQGLPFPADVFMIPQYGL----- 379
Db 304 PPTKIYNNDLTASLWALDNTNLTPTTPAAPSSETLGFYPMKPTKPYRYILSCIRNLN 363
QY 380 --TLNNGSQAVGRS-----FYCLE-YFPSSQMLRTGNNE-TFSYTFEEVPFHSSYA 426
Db 364 PPTYGSOQPNRNLNTRNLHSDIMEFYTIENAVPIHLRLTGDESFSTGIYHFDTKPL--KLT 421
QY 427 HGOSLDR-----LMNPLI--DQYLYLNRTQNGSGAQNKDLFSRGSAPGWSVQPK 476
Db 422 HSWQNTRSGLPRLKLTETPTBGDHPGLPAANTKKGHNQITNSYTEAT---AIRP- 476
QY 477 NMLPGCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINP--GTAMASHKDEDEK 534
Db 477 -----AQGVNTPYNNFEYSNGRP-----LPIPTADTYQNDDEPN- 514
QY 535 FPMGSMITGKSAGASNTALD--NWMITDEEIK--ATNP-----VATERFGTVA-V 582
Db 515 -----GAIRFTMDYQHGHLTTSQSELRFTYFNPQSKGRAPKQGFNOAPL 560
QY 583 NFOSS-----TDPATG--DVHAMGALP-----GMWQDRDYVLOGPIWAKIPIHTD 627
Db 561 NLENTNNGTLTLPDDPLGKSNMHEFNTLTNTYGPRLTALNTAFVENGQIWDKLDLDLKP 620
QY 628 HFPSPILMGFGFLK-NPPQILIKNTVPVANPPAEFSA-TKFASTQYSTGVSEVIEIW 685
Db 621 RLH---VTAPFVCKNPPQGLPVKINP---NLTDDBNADSPOQPRIITTSNPFMWKGTLT 674
QY 686 ELQKENSKRMPDEVQYTSNYAKSANVDTVUNNGLYTEPRPIG 728
Db 675 TAKMRSSNMNPIQCHTT-----TAENIGKYI-PTNIG 706

RESULT 11
VCPVIM
coat protein VP1 - minute virus of mice (strain MW1)
C:Species: minute virus of mice, murine parvovirus
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C/Accession: B23008; B29510
R:Schli, R.; McMaster, G.K.; Hilt, B.
Nucleic Acids Res. 13, 3617-3633, 1985
A>Title: DNA sequence comparison between two tissue-specific variants of the autonomous
A:Reference number: A23008; MUID:85242059; PMID:3855242
A:Molecule type: DNA
A:Residues: 1-718 <SAH>
A:Cross-references: UNIPROT:P07302; UNIPARC:UP10000174963; EMBL:X02481; NID:960918; PID:
R:Astell, C.R.; Gardiner, E.M.; Tattersall, P.
J. Virol. 57, 656-669, 1986
A>Title: DNA sequence of the lymphotropic variant of minute virus of mice, MW(i), and

A:Reference number: A29510; MUID:86115415; PMID:3502703
A:Accession: B29510
A:Molecule type: DNA
A:Residues: 1-143, 'A', 145-718 <AST>
A:Cross-references: UNIPARC:UPI0000127D6C; EMBL:ML12032
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 10.6%; Score 424.5; DB 1; Length 718;
Best Local Similarity 23.7%; Pred. No. 9e-20;
Matches 184; Conservative 104; Mismatches 299; Indels 189; Gaps 31;

```
QY 45 LVLPGYKTLGPPNGLDKGEFVNADAAALBHDKAYDQOLKAGDNPYLRYNHADAEFQERL 104
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MHPGKYKTLGPGNSLDGEPNPSDAAKHEDEAVDQYKSGKPYLYFSADQRFIDQ 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 105 QEDTSFGNLRGAVFOAKKRLLEPLGLVEGAKTAPGKKRVBQSPQBPSSSGIGKTG 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 KDAKDWGKGVGHYFFRTYRAAPKLTDS-----PG-----TSVSRAGK 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 164 -----QAPAKKRLNFG-----QTGDSSEVPDPQLG-----BPPATPAVGPPTM 203
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 RTREPAYIFINQAAKAKKLTSSAAQSSQTMDSGTSGDGNVHSAARVERADGP-- 159
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 204 ASGGGAPMADNNEGADVGNAAGNWHGDS--TWLGD--RVITSTRTWALPTYNNHLKY 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 160 GSGGG-----GSGGGGVGVSTGSDYDQTHYRFLGDGVEITALTRLVHLNMPKSENYC 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 259 QI-----SSASGASNDNHVFGYSTPMPGFFDFNRFHCHSPRDMQMLNNMGFRKR 311
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 RIRVHNTTDTISVKNMAKDDHBEQIWTPEWSLVDAANAKWMLQPSDWQICTMWSQLNVS 273
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 312 LNFELFNIOVEVLTND-----GVTTIANLITSTVOVESSEYQLPYVLGSAHOGCLPFP 367
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 LDGIFNVVLTVEQDSGGAIKIYNNDLTACMMVAVDSNNILPYTPAASMETLGEYF 333
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 368 ADVFMILQYGY-----LTLAN-----GSOAVGRSSFYCLEYFPS--OMLRTG 407
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 WKPTIASPRYRYFCVDRDLSTVENOBGTIEHNVMGTPKGMNSQFTIENTQOITLRTG 393
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 408 NNF--TFSYTFEEVPHFSSVAHSQSLDRLMN--PLIDQLYLYLNRTQNGSQAOKDLIFS 464
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 DEFTGTGYFDPTNV--KLHTWQTNQQLGQPLSTF--PRADTDAGT-----LTA 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 465 RGSFAGNSVOPKMWL-----PPCYROORVSTKTDDNNNSN 500
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 QGSHGHTOMEVNVVSEAIRTPRAOVGRCOPHDFEASRAGP--FAAPRVADVTQGVDR 500
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 501 PTWTGASKYNLNGRESIINPGTAMASHKDEDKF-----PPMSGVMI FGESAGAS 551
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 501 ANGSVRSYSGKHGEMNAHGPAPERVTWDETNGSGSDTRDGFQSPALVPPPLNGIL 560
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 552 NTALDNWMTDEEIKATNPVATERFGTVANFOS--SSTDPAQCDVHAMALPGMWOD 609
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 561 TNA-----NPLGTGN--DIHFSNVFSGYPLAFSH-----P 590
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 610 RDVYLQGPIMAK--IPHTDGHFSPPLMGFGKPNPQILIKNTP--VEANPAEESA 664
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 591 SPVYPQGIWMDKELDLHKKPRLHTAPFV--CGNNAFGQMLVRLGRLTDQYVPMNGATL 647
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 665 TKFASFTQYSTGVSVIEIWELOKENSKRNPVQTSNAYKASANDFTYDNGG 719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 648 SRIVTYGTFPMWKGKLTIRAKLRA--NTWNVPYQ-----VSVEDNG 686
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 12

vcpcvp
coat protein VP1 - canine parvovirus (strain N)
N:Contains: coat protein VP2
C:Species: canine parvovirus, CPV
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
R:Reed, A.P.; Jones, E.V.; Miller, T.J.

U:Viro1. 62, 266-276, 1988
A:Title: Nucleotide sequence and genome organization of canine parvovirus.
A:Accession number: A29962; MUID:88062992; PMID:2824850
A:Accession: B29962
A:Molecule type: DNA
A:Residues: 1-748 <RE>
A:Cross-references: UNIPROT:P12930; UNIPARC:UPI0000127D7B; EMBL:M19296
C:Genetics: 26/3
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein
P:165-748/Product: coat protein VP2 #status predicted <VP2>

Query Match 10.6%; Score 421.5; DB 1; Length 748;
Best Local Similarity 23.0%; Pred. No. 1.5e-19;
Matches 184; Conservative 114; Mismatches 287; Indels 215; Gaps 34;

```
QY 44 GLVLPGYKTLGPPNGLDKGEFVNADAAALBHDKAYDQOLKAGDNPYLRYNHADAEFQER 103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32 GLVPPGYKTLGPGNSLDGEPNPSDAAKHEDEAVDQYKSGKPYLYFSADQRFIDQ 91
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 104 LQEDTSFGNLRGAVFOAKKRLLEPLGLVEGAKTAPGKKRVBQSPQBPSSSGIGKTG 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 92 TKDAKDWGKGIHGFRAKKAIAVLTDTDPHPBSTRTK--PTKRSKRPPIHFIULAK-- 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 164 QAPAKKRLNFGQTDSESV-----DQPLGPPA--TPAAVGPPTMASGGA PMADNNEG 217
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 149 -----KKKAGAGQVRNDLAPWSDGAVQPDGQPAVRNBRATGSGNGSGGG-----GGG 198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 218 ADGVGNASGNWH-----CSTWLGDREVITTSRTWALPTYNNHLKYQI-----SSA 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 SGGVGISTGTFFNNQTEFKLENGWV--EITANSSRLVHLNMPKSENYRVVNNMDKAV 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 264 STGASNDNHVFGYSTPMPGFFDFNRFHCHSPRDMQMLNNMGFRKLNFLFIQVKE 323
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 NGMALDDIHAQIYTPMSLVDAANAKWGFENGDQILVNTMSELHVSFEOEIFNVVLKT 316
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 324 VT--TNDGVTTIANLITSTVOVESSEYQLPYVLGSAHOGCL-----PPFPA----- 368
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 VSESATQPTKVYNNDLTASLMLVALDSNNTMPFPAAWRSBTLGFPYKPTIPFPWRYF 376
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 369 --DVFMILQYGYLTNLNGSAVGRS-----SFCLE--YFSPQMLRTGNF--TFSY 414
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 QMDRTLLPSH-----TGTSGTPNIYHGTDPPDVQFYTINSVPVHLRTGDEPATGTF 430
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 415 TFEVPHFSSVAHSQSLDRLMN--PLIDQLYLYLNRTQNGSQAOKDLFSRGSFAGMS 472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 431 FFDCKP--CRLTHTWQTNRALGLP-----FLNSLPQSGATNFGDI-----GV- 472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 473 VOPKMWLPGCYROORVSTKTDDNNNSN-----TWTGASKYNLNGRESIINP--GT 522
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 473 -----QDDKRGVTOGMNTNYITLTIATIMPAEYGYAPYSPASQGPRTK 519
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 523 AMASHK-----DDEDKFPMSGVMI FGESAGASNTALD----- 556
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 520 PIAARGGAQTYENQADGDPY-----AFGQHQOKTITTBETBERFTYIAHODTGR 572
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 557 -----NVMITDEEIKATNPVATERFGTVANFOSSTDPAQCDVHAMALPG 604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 573 YPBGDWQININFLPVINDVNLFTDPIG--SKTGINY--TWIPNTYGLTALANVP- 625
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 605 MWQDRDVLQGPIMAKIPHTDGHFSPPLMGFGK-----NPPQILIKN 651
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 626 -----PYVFNQGIWMDKEPDT-----LKRPLHVNAPFVCONNCPQLPFVKV 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 652 TPVPAN--PPAESAKFASFTQYSTGVSVIEIWELOKENSKRNPVQTSNAYKS 708
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 667 APNLTNEYDPPDASANMR-----IVTYSDFMWKGLVFKAKLRASHATWNPIQOMSIN----- 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 709 ANVDFTVDNNGLYTEPRPIG 728
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 719 -----VDNQPNVY--PSNIG 731
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


This Page Blank (uspto)

QY 324 VTTNDGTTAANNITSTVQVPSD-----SEYQLPYVLGS-----AHQCL 363
| : : : : :
Db 402 SSTAFVGESEAMAKPTBSKKTFSNBSAGNESSTSAAPLNGSIFSGANTVTPPPNSGLT 461
| : : : : :
QY 364 - - - PEPFADVMIPQYGLTLNNGSQAVGRSSFYCLEYFPQGMALTGNFTSTYFEVPPF 421
| : : : : :
Db 462 SSBSPFTTISNIP-----SDNSVGD-----MPSTVQSPAAITHNSSIFGLKLP 504
| : : : : :
QY 422 HSSAYASQSLDRMLNPLIDQVLYLNRTQNSGSAONKDLIFSRSAG-MSVQPKWLP 480
| : : : : :
Db 505 SMD-----SNSQTSASPLSSTSPFKGQAPAPSAVAESS 542
| : : : : :
QY 481 GPCYRQORVSKTKTDNNNSNFTWTG-----ASKYNLNGRE----- 515
| : : : : :
Db 543 GQISKEDEV-KNAFTFGNTSTFKFGKMASADOSTGTSVGAESAENKSRBPFGVSSSVVG 601
| : : : : :
QY 516 SIINPGTAMASHDDEKFFPMGVMIFGESAGASTALDNWMTDEBEIKATNPVATE 575
| : : : : :
Db 602 STLNPTAALAAAPES-----SGSLIFGVTSSTPGTETSKISASS---AATN-TGNS 649
| : : : : :
QY 576 RFGTVAVNFQSS-----STDPATGVHAMGALPGMWODRDVYLQGPVAKIPIHT 625
| : : : : :
Db 650 VFGTSSAFATSSSGSMWGVASATGSSVFGFNVAVSASATSSQSAENRFGAGNATQNT 709
| : : : : :
QY 626 DG-----HFHPSPLMGFGFLKNPPOILLIKNTVPVAN-PPAEFSATKEASFTQ 673
| : : : : :
Db 710 GSGTTTSTQSIPOFGSSPAPSPGLSG-----NSSLASNSGPGFSKSEBAPV-TS 760
| : : : : :
QY 674 YSTGQVS 680
| : : : : :
Db 761 GSTPQLS 767

RESULT 4

US-10-471-571A-3542
; Sequence 3542, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927NO
; CURRENT APPLICATION NUMBER: US/10/471. 571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqMan99, version 1.03
; SEQ ID NO 3542
; LENGTH: 1629
; TYPE: PRP
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(1629)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-3542

Query Match 3.1%; Score 122.5; DB 6; Length 1629;

Best Local Similarity 20.1%; Pred. No. 0.56; Indels 251; Gaps 41;

Matches 158; Conservative 100; Mismatches 278; YQSKVKIEVY 78

QY 45 LVLPQYKYLGFNFGLDKGEFVNADAAALEHDKAYDQOLKAGDNPYLRYNHADAEPQERL 104
| : : : : :
Db 36 ILIGSLMYLG-----TQGEAEAAENNIENPTLTK--DN---YQSKVKIEVY 78
| : : : : :
QY 105 QEOTSTFGNIGRAVFAQAKKVL--EPGLVBEAGAKTAPGKKRPVQSPQEDSSSGTGKT 162
| : : : : :
Db 79 NKDTAPGVAKSEVSNKDTIEHEPVSUKADISKEDTKEVADVAVQPKKS--VTNH 136
| : : : : :
QY 163 GOOPAKKRL-----NFGOTGDSSEVPD-----PQDLGPRPA 193
| : : : : :
Db 137 AETPKYRKAASVDEGSDITRDSKNVVESTPIITIQGKEHEBGSGVDIQKKPDLGLGSEV 196

QY 194 TPAAVGPTTMAAGGAPMAD-----NNEGADG-----VGNASGN 227
| : : : : :
Db 197 TRFVNGESNGLIGALQLOKNKIDSKDFNFYKRVANNHQSNWTGADGGLFSLKGNAL-E 255
| : : : : :
QY 228 WHCSTLGDVRVITTSRTALPT--YNNHLYQOISASVGSANDNHYPGYSTPWWGFD 284
| : : : : :
Db 256 YLTNGIGLGDGLVNS--GGFKIDIGYIYTSSMDTEKQAGG-----YROY---GAF- 303
| : : : : :
QY 285 FNRPHCHFSPRDWQLINNMMGFRPKRLNFKLFIQVKEVTTNDGVTTLAANNITSTVQVF 344
| : : : : :
Db 304 -----VANDSG-----NSQWGENIDKSKTN--FLNTAANDSTNT----- 336
| : : : : :
QY 345 SDSEYQ-----LPYVLGSAHQG-----LPPFADVMIPQYGYTL 381
| : : : : :
Db 337 SDGKFGQRLNDVILTYVASTGKRAEYAGKTWETSITDGLSKQAVNPLITSORWGL 396
| : : : : :
QY 382 NNGSQAVGRSSFYCLEYFPQGMALT--GNFTTS-----YTFEEVPHSSYAH 427
| : : : : :
Db 397 NQGINANG-----MMRTDLKGSFETTPPAKPTITTELEKKVEBIPKK--- 439
| : : : : :
QY 428 SQSLDRMLNPLIDQVLYLNRTQNSG-----SAQN--KDLFSRGP-AGMSVQPKN 477
| : : : : :
Db 440 -----ERKFNPDLAGTEKAYR-EGQKEKITTPTLKNPLTGVLIISKGEPEBELTXDPIN 494
| : : : : :
QY 478 WLP-----GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINP--GTAMASHK 528
| : : : : :
Db 495 ELTEYGEETIAPGHRDEFDPLPT-----GEKEEVPKPGIKNPETGDVVRPV 543
| : : : : :
QY 529 DDEKFFPMGVMIFGESAGASTALDNWMTDEBEIKATNPVA--TERGTYAVANQS 586
| : : : : :
Db 544 DSVTKYGPVXGDSIVEKE-----IPEKEREKFPDLAPGTEKVTREGQKGEK 591
| : : : : :
QY 587 SST-----DPATGVHAMGALPGMWOD--RDVYLQGPVAKIPIHTDGHFAP-----S 632
| : : : : :
Db 592 TITPTLKNPLTGEIIGKESKEITKDPINELTEYGPETITPGRD-EPFPPKPTGKE 650
| : : : : :
QY 633 PLMGFGFLKNPPOILLIKNTVPVANPPAEFSATKEASFTQYSTGVQSVETIEMLOKENS 692
| : : : : :
Db 651 EVPGKPGIKNPETGDVVR-----PVD-SVTXYGPV-----XGDSIVEKE-EIPFEKE 696
| : : : : :
QY 693 KRNPEV 699
| : : : : :
Db 697 RKNPDL 703

RESULT 5

US-10-505-928-393
; Sequence 393, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505.928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patentin 3.2
; SEQ ID NO 393
; LENGTH: 667
; TYPE: PRP
; ORGANISM: Homo sapiens
US-10-505-928-393

Query Match 3.0%; Score 119; DB 6; Length 667;

Best Local Similarity 20.9%; Pred. No. 0.3; Indels 192; Gaps 33;

Matches 118; Conservative 75; Mismatches 179; SSSGSN-----GNGGHPSPSKNYG 72

QY 191 PPAVPAVGPPTMASG--GGAPMADNNEGADGVGNASGNHCDSTWLGADVITTSRTWA 248
| : : : : :
Db 28 PVSXGKNGPTSLASGHFTGSNVEDR-----SSSGSN-----GNGGHPSPSKNYG 72


```
QY 359 HQGLPPPADVPMIPOYGYTLNNGQAVGRSSPYCLEYPSQMLRTGNFTSYFEE 418
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

Query Match 2.6%; Score 103.5; DB 7; Length 561;
Best Local Similarity 19.6%; Pred. No. 3.2;
Matches 70; Conservative 57; Mismatches 157; Indels 73; Gaps 14;

QY 280 WGYDFRFRHCFRPMQRLINNKG--FRPKRLNFKLNIQKXYVTNNGVTIANNL 337
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 201 WG-----QDMQVMLNNAASEVAQADSLSTMTFSVHMVQVYNN--LSTVENYV 245
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 338 TSTVQVESDSERYQLPYVL---GSAHQGLPPPADVPMIPOY--GYLTLN---NGS--- 385
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 246 STFL-----SSHNLPLVGERGADHQBEVEDAILSAEQYGYGLMSGSGSCGT 300
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 386 -QAVGRSSPYCLEYPSQMLRTGNFTSYFEEVPHSYAHSQSLDRLNPLIDQYLY 444
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 301 LDTNNFNVSILTSWGRLLINGTIGIKATSVIASVYSGSSSSSSSSSSSSSSSS 348
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 445 YLNRTOQSGAOKDILLFSGRSPAGMSVOPKMLPGPCY-----RQGRVS 490
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 349 --SSTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSG 403
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 491 KTKTDNNNSNFTMGASKYKYNLNGRESIINPGTAMASHKDDKFFPMGVMIFGKESAGA 550
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 404 RTTESGSGNGGVIGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSG 463
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 551 SNTLDNVMITDEBEIRATNPVATERGCTVAVNFQSSSTDPATGDVHAGMLPGMNV 607
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

QY 464 SGATCEHI-ITNSWNSGFOGAVRITNNGSSAINGVQVSWSS-----YSDGTTISVW 513
; Sequence 145, Application US/11211154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKBORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 561
; TYPE: PR1
; ORGANISM: Microbulbifer degradans
US-11-121-154-145

RESULT 9
US-11-293-697-4614
```

```
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-293-697-4614

Query Match 2.6%; Score 102.5; DB 7; Length 684;
Best Local Similarity 25.9%; Pred. No. 5;
Matches 68; Conservative 28; Mismatches 90; Indels 77; Gaps 15;

QY 26 KCGAPKPKRANQKDDGRGLVLPGYKYLGPENGLDKG-EPVNAADAALBHDKAYDQ-QL 83
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 156 KKPDPKPKVKKLY-----HQYIPDQKAEKPPWDASAYARLQQQLFLQQLI 205
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 84 KAGDPYLRYNHADAERQERLQEDTSGNIGRAVFOAKKRVLEPLGLVEGAKTAPGKK 143
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 206 L-----SQQQQQQHRSYLG-----MQAQ-----LX 228
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 144 RPEVQSGFOEPDSS-GIGKTGQOPAKKRLNFGQGDSESVDDPOPLGEPATPAVGGTT 202
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 229 EPNEQWNRNPNSSSTPLSNTPLEPVKNFS-GQTGVSSFKGPPLP---PNLDLKVSELR 284
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 203 M-----ASGGAPADN-NEGADGVNASGNMHODSTWLGPRVITTSRTVALPTYN 253
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 285 QQLRIRGLPVSGYTALMDRLRPQDCSGNPNVNF-----GD-ITVT---FVTP 331
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 254 NHLVKQISASTGA-SNDNHYFG 275
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-293-697-4614

QY 332 NTLFNYQSSSTSLNSNGFYHFG 354
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-293-697-4614

RESULT 10
US-10-985-570-2

Query Match 2.6%; Score 102.5; DB 6; Length 1464;
Best Local Similarity 23.4%; Pred. No. 15;
Matches 57; Conservative 6; Mismatches 62; Indels 119; Gaps 9;

QY 26 KPGAPKPKRANQKDDGRGLV---LPGYKYLGPENGLD-----KGEVNAADA 70
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-985-570-2

QY 237 KPGRPGRGPGPGP-GARGLPGTAGLPKMKGNRFGSGLDGAKDAGPAGPGEPSGNGEN 295
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-985-570-2

QY 71 AALEHDAKAYDQQLKAGDPYLRYNHADAERQERLQEDTSGNIGRAVFOAKKRVLEPLG 130
; Sequence 4614, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4614
; LENGTH: 684
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-985-570-2
```


Db 296 GA-----PGMG-----PRG 305
QY 131 LV-EEGAKTAPGKKRPVEQSPQEPDSSSGTGTGQOPAKRLINFGQTGSESVDPDPLG 189
Db 306 LPERGRPGAPG-----PAGARGNDGATG--AAGPFGTGP 340
QY 190 EPPATPAV-----GPTMASGGA PMADNNEGADGVG 222
Db 341 GPPGFGAAGVAGKEAGPQGPGRSGPGVARGEPGPAGAPAGNPGADGQPGAKGAN 400
QY 223 MASG 226
Db 401 GAPG 404
RESULT 11
US-11-246-999-145
; Sequence 145, Application US/11246999
; Publication No. US20060099622A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/11/246,999
; PRIOR FILING DATE: 2005-10-11
; PRIOR APPLICATION NUMBER: US/09/984,130
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 145
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-246-999-145
Query Match 2.5%; Score 101.5; DB 7; Length 439;
Best Local Similarity 20.3%; Pred. No. 3.2;
Matches 71; Conservative 49; Mismatches 90; Indels 129; Gaps 20;
QY 293 SPEDMQRLLNNMGFRPKR-----INFKLFNQVKEVTNNGV-----TTIAN 335
Db 53 NPGGFTFLDNHWGLDQRYRLIHYSEDELSTIRLSNTVHD---EGVYCYYSSTPFRS 108
QY 336 NLTS-----TVQVFSDESEYQLPYVLGSAHQGLPPPADVFMIPOYGYLLTNGS 385
Db 109 KMTTVEVLAPSKRPVLOVSRDTBGRV--TLSCYTQCKPQ-----PQITML-LDNGI 157
QY 386 QAVG-----RSSFYCLEYPPSQMLRT-----GNNFTFSYTFEEVPRH 422
Db 158 QLTGDRHKLKLEADGKWTSTLTVLVAYGPNSTATCLVHHKALGGKLTPEPQFEDVA-- 215
QY 423 SSVAHQSGLDRMLNPLIDVYL-VYLNRTQNSGSAQKDLLFSRGSAPAGMSVQPKMLPG 481
Db 216 RYVANTTPVSTTLE--VDIVSEVVOPTVTTASDLNSTDPS----- 256
QY 482 PCYRQORVSKTKTDNNNSFT-----WTGASKYNL-NGRESIINPTAMASHKDEDKP 534
Db 257 PSIFPQ-----HNSGATSVAAAGELSGTSAHHIEGTETALN-CTV-----TEELF 300
QY 535 -----FPMGVMIFGKESAGASNTALDNVMTDEETIKA 568
Db 301 RTEASFSEENVTLIS-----IVTFEODVKS 325

RESULT 12
US-10-505-928-759
; Sequence 759, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 759
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-759
Query Match 2.5%; Score 101.5; DB 6; Length 1049;
Best Local Similarity 18.2%; Pred. No. 11;
Matches 122; Conservative 79; Mismatches 179; Indels 291; Gaps 35;
QY 122 KKRVLRLPLGVEGA-KTAPGKKRPVEQSPQEPDS-----SSGIGKTGOOPAK-KRLN 172
Db 82 KANTSOP-GVLQGAAYLLCPWGAASPTQCTPIEDFSKGRLESLSSEGEVEYKSLQ 140
QY 173 -FGQT-----GDSESVDP-----QPLGEPPATPAVGPPTMASGG-----GAPMD 213
Db 141 WEGATVRAHGSSTIACAPLYSWRTEKEPLSDP-----VGTCLSTDNTRLILEVAPCRS 194
QY 214 NNEGADGVGNASNMHCDS-----WIGDKVITSTITMALPTN-----N 254
Db 195 DPSMAAGGCGCGGFAEFTKGRVVLGGPGSYFWGQ--LISATQEIJAESTYPEYLIN 252
QY 255 HLYKOISSASTGASNNHYFGYSTPMGYFDENRHFCHFSPRDQRLNNMGFRPYRLNF 314
Db 253 LVQGGQLQIROASSIYDSTLIGISVAVG-----EFGSD----- 285
QY 315 KLFNIQKEVTNNDGVTTI-ANNLSTVQVFSDESEYQLPYVLGSAHQGLPPPADVFMI 373
Db 286 -----TEDFVAGVPKGNLT----- 299
QY 374 POYGYLLTNGSGAVRSSFYCLEYPPSQMLRTGNNFTFSYTFEEVPRHSSVA-----H 427
Db 300 --YGYVTIINGSD-----IRS-----LYNFSGEOMASYFGYVAATDVN 336
QY 428 SQSLDRLM--NPLIDQYLYLNRTQNSGSAQKDLLFSRGSAPAGMSVQPKMLPGPCYR 485
Db 337 GGDGLDLVLGAPFL-----MDRTPPGRPOEVGRVYVYLQ-HRAGIEPFP----- 379
QY 486 OGRVSKTKTDNNNSNFTWTGASKYKYNLNGRESIINPTAMASHKD-DEKFFPMGVMIFG 544
Db 380 -----TLTLAGHDEF--GRF-----GSSILPLDLDLDDGDNDVAIGAIFG 417
QY 545 KESAGASNTALDNVMTDEBEIKATNPATERGTYAVNFQSSSTPAPAGDVAMAGALPG 604
Db 418 GEF-----OQGVVFN-----FPGGPGLSKPS 440
QY 605 MWQODRDVYLQGIWAKIPHT-----DGHFSPPLMGSGFGKLNPPOLIK 650
Db 441 QVLQ-----PLMA-ASHTPDFGSALRGGDDGNGYFDLIVGSFGVDK--AVVYR 488
QY 651 NTPV-----PA-----NPPA-----EFSATKFAFTIQTSGOVSVEI 683
Db 489 GRPIVASASLITFPMMPNEERSCSLEGVPVACINLSFCINAGKHAVDSIGFTVELQ 548
QY 684 EWELOKENSKR 694
Db 549 DMQOKGQGVRR 559

```
RESULT 13
US-11-121-154-172
; Sequence 172, Application US/11121154
; Publication No. US20060105914A1
; GENERAL INFORMATION:
; APPLICANT: TAYLOR, LARRY EDMUND
; APPLICANT: WEINER, RONALD M.
; APPLICANT: HUTCHESON, STEVEN WAYNE
; APPLICANT: EKEORG, NATHAN A.
; APPLICANT: HOWARD, MICHAEL
; TITLE OF INVENTION: PLANT WALL DEGRADATIVE COMPOUNDS AND SYSTEMS
; FILE REFERENCE: 108172-00121
; CURRENT APPLICATION NUMBER: US/11/121,154
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,971
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 172
; LENGTH: 1186
; TYPE: PRT
; ORGANISM: Microbulbifer degradans
US-11-121-154-172

Query Match
Best Local Similarity 19.8%; Pred. No. 13;
Matches 85; Conservative 50; Mismatches 139; Indels 155; Gaps 21,

QY 349 YQAPYVIGSAMHQLGPPF-----PADVPMTPQVGYLT-----LNG 384
DB 211 YRTQVVGGAASVPTPHKMGVSFLGEMKIGDPAGA-----GYTPDDITARIKNG 264
QY 385 SOAVGRSSFYCLEFFPSQMLRTGNNFTFSY--TFEEVPHSSVYASOSLDRIMNPLIDQY 442
DB 265 VRIMG-----IPNLGAGNQGYTSVDPDPSEVVDGIAVANSEYAN-----LEAY 309
QY 443 L-YLLNRT--QNGSGAQNKKLLFSRGSFAGMSVQPKNMLPGCYRQORYSKTKTDNNN 498
DB 310 LKSYSDGTATVQMOQGNLPVMQATFVHGSP-----YVFPAKAYRGNMVLRTKADCG 360
QY 499 SNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSVMTPGKESASASTALDNV 558
DB 361 EKGTF--YNNNSLGI--WTSVAGNKNQD-----FLITGEGEVFNNIETDTI 403
QY 559 MITD-----BEEKATNPVA-----TERFGTVAV 582
DB 404 TLTNAANEFTLTLLPTAGACTPSSSTVIAQFEDSARAV--VAKDIOYSVDXTNNMVTYTH 461
QY 583 NFOSSSTDPAATGDVHANGALPGWVQODRDVYLQ-----PIWAKIPHTDGHFPPSLMG 636
DB 462 TYKNESNTP--YQTLAGLLPMHMKYSDTLGKYKTRSGARGMVQFAHIDFSYTIPIYVG 517
QY 637 GFGLAKNPPOLLIKNTVPANPRPASFATKRFASFTIQTQSVQVELEMELOKNSKRN 696
DB 518 -----VL-----PYLPSVGDFFDSVLAGLVQAFVA-----SGPENMN 550
QY 697 PEVOYTSNY 705
DB 551 P--HTDTY 556

RESULT 14
US-10-953-349-23198
; Sequence 23198, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
```

```
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23198
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-23198

Query Match
Best Local Similarity 19.8%; Pred. No. 3.6;
Matches 63; Conservative 21; Mismatches 97; Indels 137; Gaps 14;

QY 5 GYLPDWLEDNLSGIREMWDLKPGAPKPKANQKODDGRGLVLPQYKYLQ--PENGLDKG 62
DB 131 GYPP-----QQPDYSTWMDOSTAP--HQSTHGG-----GYVYSQOPQOPQPNPG 175
QY 63 EPVNAADAALAEHDK-----AYDQQLKGG--DNPTLRVNHADAEPERLOEDPSSFGNIGR 116
DB 219 GPAPPADGSAYNYSQPPSSGYNQSAQGYAODSTNAYN-----AQSGSGYG----- 264
QY 117 AVFOAKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGOQPAKK--RLNFG 174
DB 265 -----GPTVDDQGGYG--SSNPAQEGHTANYA 291
QY 175 QTGDSSEVPDPPL-----GEPPATPAVAGPTTMA 204

RESULT 15
US-10-953-349-23197
; Sequence 23197, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23197
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-23197

Query Match
Best Local Similarity 19.8%; Pred. No. 4.2;
Matches 63; Conservative 21; Mismatches 97; Indels 137; Gaps 14;

QY 5 GYLPDWLEDNLSGIREMWDLKPGAPKPKANQKODDGRGLVLPQYKYLQ--PENGLDKG 62
DB 174 GYPP-----QQPDYSTWMDOSTAP--HQSTHGG-----GYVYSQOPQOPQPNPG 218
QY 63 EPVNAADAALAEHDK-----AYDQQLKGG--DNPTLRVNHADAEPERLOEDPSSFGNIGR 116
DB 219 GPAPPADGSAYNYSQPPSSGYNQSAQGYAODSTNAYN-----AQSGSGYG----- 264
QY 117 AVFOAKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGOQPAKK--RLNFG 174
DB 265 -----GPTVDDQGGYG--SSNPAQEGHTANYA 291
QY 175 QTGDSSEVPDPPL-----GEPPATPAVAGPTTMA 204
```

Db	292	VQGDShQAPSAQPTTVAAQOGYPTNQLPSSNTANYPPQGTTPQPGYGVPTSGAAVGMQSO	351
Oy	205	SGGGA PMADNNE-----GADGVGNASGNMHCDSTWLGDRVITSTRWALPTYNHL	256
Db	352	GYGGAVQPGYGPQTYGAPQGGGPGYQ-----ALPSYSNS	387
Oy	257	Y---KOISSASTGASNDN	271
Db	388	YGAGYTQTPAYTGDNGN	405

Search completed: June 12, 2006, 12:57:44
 Job time : 19 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2006, 12:43:59 ; Search time 305 Seconds
(without alignments)
2232.169 Million cell updates/sec

Title: US-10-696-261-13

Sequence: 1 MADDGYLPMLEDNLESGIR.....NNGLYTEPRDITGRYLTPRL 736

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3989	100.0	736	Q9WB88_9VIRU	Q9WB88 adeno-ssoc
2	3968	99.5	736	Q6JC08_9VIRU	Q6JC08 adeno-ssoc
3	3963	99.3	736	Q56137_9VIRU	Q56137 adeno-ssoc
4	3883.5	97.4	737	Q6JC13_9VIRU	Q6JC13 adeno-ssoc
5	3873	97.1	736	Q6JC12_9VIRU	Q6JC12 adeno-ssoc
6	3853	96.6	736	Q6JC10_9VIRU	Q6JC10 adeno-ssoc
7	3511	88.0	736	Q56139_9VIRU	Q56139 adeno-ssoc
8	3494	87.6	736	Q65311_9VIRU	Q65311 adeno-ssoc
9	3473	87.1	736	Q808X3_9VIRU	Q808X3 non-human P
10	3469.5	87.0	735	Q67008_9VIRU	Q67008 adeno-ssoc
11	3456.5	86.7	735	Q6JBE2_9VIRU	Q6JBE2 adeno-ssoc
12	3456.5	86.7	737	Q6JC52_9VIRU	Q6JC52 adeno-ssoc
13	3455.5	86.6	737	Q6JC58_9VIRU	Q6JC58 adeno-ssoc
14	3453.5	86.6	735	Q6JC28_9VIRU	Q6JC28 adeno-ssoc
15	3453	86.6	738	Q6JC19_9VIRU	Q6JC19 adeno-ssoc
16	3451.5	86.5	735	Q6JC34_9VIRU	Q6JC34 adeno-ssoc
17	3450.5	86.5	737	Q6JC51_9VIRU	Q6JC51 adeno-ssoc
18	3449.5	86.5	735	Q6JBE1_9VIRU	Q6JBE1 adeno-ssoc
19	3449.5	86.5	735	Q6JC42_9VIRU	Q6JC42 adeno-ssoc
20	3449	86.5	738	Q5Y9B4_9VIRU	Q5Y9B4 adeno-ssoc
21	3449	86.5	738	Q6JC37_9VIRU	Q6JC37 adeno-ssoc
22	3448	86.4	738	Q808W5_9VIRU	Q808W5 non-human P
23	3446.5	86.4	735	Q6JC17_9VIRU	Q6JC17 adeno-ssoc
24	3445	86.4	735	Q6JC60_9VIRU	Q6JC60 adeno-ssoc
25	3443.5	86.3	737	Q6JC46_9VIRU	Q6JC46 adeno-ssoc
26	3442.5	86.3	735	Q6JC36_9VIRU	Q6JC36 adeno-ssoc
27	3442.5	86.3	737	Q8JQ60_9VIRU	Q8JQ60 adeno-ssoc
28	3442	86.3	738	Q6JC14_9VIRU	Q6JC14 adeno-ssoc
29	3442	86.3	738	Q6JC56_9VIRU	Q6JC56 adeno-ssoc
30	3441.5	86.3	735	Q6JC44_9VIRU	Q6JC44 adeno-ssoc
31	3440	86.2	738	Q6JC15_9VIRU	Q6JC15 adeno-ssoc

32	3440	86.2	738	2	Q6JC16_9VIRU	Q6JC16 adeno-ssoc
33	3440	86.2	738	2	Q6JC62_9VIRU	Q6JC62 adeno-ssoc
34	3439	86.2	738	2	Q6JBE2_9VIRU	Q6JBE2 adeno-ssoc
35	3438	86.2	734	2	Q6JC02_9VIRU	Q6JC02 adeno-ssoc
36	3438	86.2	738	2	Q808X3_9VIRU	Q808X3 non-human P
37	3437	86.2	738	2	Q6JC47_9VIRU	Q6JC47 adeno-ssoc
38	3437	86.2	738	2	Q6JC61_9VIRU	Q6JC61 adeno-ssoc
39	3436	86.1	734	2	Q6JC04_9VIRU	Q6JC04 adeno-ssoc
40	3434.5	86.1	735	2	Q6JBE6_9VIRU	Q6JBE6 adeno-ssoc
41	3434	86.1	728	2	Q808X4_9VIRU	Q808X4 non-human P
42	3434	86.1	738	2	Q6JBE3_9VIRU	Q6JBE3 adeno-ssoc
43	3433	86.1	738	2	Q6JC49_9VIRU	Q6JC49 adeno-ssoc
44	3433	86.1	738	2	Q6JC53_9VIRU	Q6JC53 adeno-ssoc
45	3432	86.0	738	2	Q6JC54_9VIRU	Q6JC54 adeno-ssoc

ALIGNMENTS

RESULT 1									
Q9WB88_9VIRU	PRELIMINARY;	PRT;	736 AA.						
ID Q9WB88_9VIRU									
AC Q9WB88_9VIRU									
DT 01-NOV-1999, integrated into UniProtKB/TrEMBL.									
DT 01-NOV-1999, sequence version 1.									
DT 07-FEB-2006, entry version 18.									
DE Capsid protein.									
OS Adeno-associated virus - 1.									
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.									
OX NCBI_TaxID=85106;									
RN [1]									
RP NUCLEOTIDE SEQUENCE.									
RX MEDLINE=99214338; PubMed=10196295;									
RA Xiao W., Chirmule N., Berta S.C., McCullough B., Gao G., Wilson J.M.;									
RT "Gene therapy vectors based on adeno-associated virus type 1."									
RL J. Virol. 73:13994-4003(1999).									
CC -----									
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms									
CC Distributed under the Creative Commons Attribution-NonDerivs License									
CC -----									
DR EMBL: AF063497; AAD27757.1; -; Genomic DNA.									
DR SMR: Q9WB88; 217-736.									
DR GO: GO:0019028; C:viral capsid; IEA.									
DR GO: GO:0005198; F:structural molecule activity; IEA.									
DR InterPro: IPR001403; Parvo_coat.									
DR Pfam: PF00740; Parvo_coat; 1.									
SQ SEQUENCE 736 AA; 81375 MW; CFABFB9BD5CD0595 CRC64;									
Query Match 100.0%; Score 3989; DB 2; Length 736;									
Best Local Similarity 100.0%; Pred. No. 4e-251;									
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy 1	MADGYLPMLEDNLESGIREWMDLKGAPKPKANQKODGRGLVPGYKYLGPFGGLD	60							
Db 1	MADGYLPMLEDNLESGIREWMDLKGAPKPKANQKODGRGLVPGYKYLGPFGGLD	60							
Qy 61	KGEPVNAADAALAEHDKAYVQOLKAGDNPYLRYNHADAERQERLQEDTSPFGNIGRAVFQ	120							
Db 61	KGEPVNAADAALAEHDKAYVQOLKAGDNPYLRYNHADAERQERLQEDTSPFGNIGRAVFQ	120							
Qy 121	AKKRVLEPLGLVEGATAPGKKRPVQSGFQEPDSSGIGKTGQPPAKKRLNFCQTGDS	180							
Db 121	AKKRVLEPLGLVEGATAPGKKRPVQSGFQEPDSSGIGKTGQPPAKKRLNFCQTGDS	180							
Qy 181	SVDPPOPGLGEPATPAVGPPTMASGGGAPMADNNEGADGVGNAGMHCSTYLGDRV	240							
Db 181	SVDPPOPGLGEPATPAVGPPTMASGGGAPMADNNEGADGVGNAGMHCSTYLGDRV	240							
Qy 241	TTSTRITWALPTYNHLYKQISSASTGASNDNHFGYSTPWCYEDFNHFGHFSFRDQRL	300							
Db 241	TTSTRITWALPTYNHLYKQISSASTGASNDNHFGYSTPWCYEDFNHFGHFSFRDQRL	300							
Qy 301	INNMGPRPKRLNFKLFNIQVKEVTTNDGVTTIANLITIVQVPSDSEYQLPYLSAHO	360							

Db	301	INNNGFRKRLNFKLNFLENQVKEVTTNDVTTIANNLSTVOQFSDSEYQLPYVLGSAAHQ	360
Qy	361	GCLPFPADVPMIPOYGYLTLLNNGSOAVGRSSFYCLEYFPSCWLRTGNNFTFSYEEYBP	420
Db	361	GCLPFPADVPMIPOYGYLTLLNNGSOAVGRSSFYCLEYFPSCWLRTGNNFTFSYEEYBP	420
Qy	421	FHSYAHSGSLDRMLNPLIDQYLYLNRITONOGSGAQNKDLEFSRGSAGMSYQPNWLP	480
Db	421	FHSYAHSGSLDRMLNPLIDQYLYLNRITONOGSGAQNKDLEFSRGSAGMSYQPNWLP	480
Qy	481	GPCYHQQRYSKTKYTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMISGY	540
Db	481	GPCYHQQRYSKTKYTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMISGY	540
Qy	541	MIFGKESAGASNTALDNWMITDBEEIKATNPVATBPGYAVAVFOSSSTDPAQGDVHAHG	600
Db	541	MIFGKESAGASNTALDNWMITDBEEIKATNPVATBPGYAVAVFOSSSTDPAQGDVHAHG	600
Qy	601	ALPGVWVQMODRDVYLQGPIMAKIPTHDTGHFHPSPIMGSGFGLKNPPQILLKNTFVPANPRA	660
Db	601	ALPGVWVQMODRDVYLQGPIMAKIPTHDTGHFHPSPIMGSGFGLKNPPQILLKNTFVPANPRA	660
Qy	661	EFSSATKFASTFOYSTGQVSVIEIWELOKENSKRANPEVOYTSNYAKSANVDFTVNNGL	720
Db	661	EFSSATKFASTFOYSTGQVSVIEIWELOKENSKRANPEVOYTSNYAKSANVDFTVNNGL	720
Qy	721	YTEPRPIGRTYLRPL	736
Db	721	YTEPRPIGRTYLRPL	736

```

RESULT 2
ID Q6JC08_V9IRU PRELIMINARY; PRT; 736 AA.
AC Q6JC08;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Capsid protein VP1.
GN Name=caps;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
OC unclassified Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed:15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenbergh L.H., Alviria M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AY530611; AAS92926.1; -; Genomic_DNA.
DR SMR: Q6JC08; 217-736.
DR GO: GO:0019028; C:Viral capsid; IEA.
DR GO: GO:0005198; F:Structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo coat.
DR Pfam: PF00740; Parvo coat_1.
SQ
SEQUENCE 736 AA; 81317 MW; 37307C40975CDD9D CRC64;

Query Match 99.5%; Score 3968; DB 2; Length 736;
Best Local Similarity 99.3%; Pred. No. 9,4e-250;
Matches 731; Conservative 3; Mismatches 2; Indels 0; Gaps

QY 1 MAADGYLPDWLIEDNLSEGIREWMDLKGAPKPKXANQKQDGDGGLVPGYKTYGPPNGSD 60
DB 1 MAADGYLPDWLIEDNLSEGIREWMDLKGAPKPKXANQKQDGDGGLVPGYKTYGPPNGSD 60

```

QY	61	KGEPRNADDAALAEHDKAYDOOLKAGDNPYLRYNNADEAFORLEJEDTSFEGULGAVFQ	120
Db	61	KGEPRNADDAALAEHDKAYDOOLKAGDNPYLRYNNADEAFORLEJEDTSFEGULGAVFQ	120
QY	121	AKKRVLBEPLGVEEESAKTAPGKKRPEVOSPOEBSSSGIGTKQOQPAKKRLNFQGTGDE	180
Db	121	AKKRVLBEPLGVEEESAKTAPGKKRPEVOSPOEBSSSGIGTKQOQPAKKRLNFQGTGDE	180
QY	181	SVPDPQJGEPAPATAAVPTTMAAGGGA PMADNNEGADGVCNASENHHCDSFTWLGDRYI	240
Db	181	SVPDPQJGEPAPATAAVPTTMAAGGGA PMADNNEGADGVCNASENHHCDSFTWLGDRYI	240
QY	241	TTSTRTMALPTYNMLYKOISSASTGASNDNHYFGYSTPMGYFENRFPHCSFPRMORL	300
Db	241	TTSTRTMALPTYNMLYKOISSASTGASNDNHYFGYSTPMGYFENRFPHCSFPRMORL	300
QY	301	INNMGPPPKRLNFKL FNI OVEKEVTITNDGVTITANNLSTVOYFSDSEYQLYVLGSAHQ	360
Db	301	INNMGPPPKRLNFKL FNI OVEKEVTITNDGVTITANNLSTVOYFSDSEYQLYVLGSAHQ	360
QY	361	GCLPFPFADVFMI POXYGLYTLNNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSEYEEVY	420
Db	361	GCLPFPFADVFMI POXYGLYTLNNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSEYEEVY	420
QY	421	FHSSVAHSQSIDLRLNPLIDQYLYLYLNRTONOGSGAQNKDLLFSRGS PAMGMSVOYKOWLP	480
Db	421	FHSSVAHSQSIDLRLNPLIDQYLYLYLNRTONOGSGAQNKDLLFSRGS PAMGMSVOYKOWLP	480
QY	481	GPCYQOQVSKTKTDNNNSNFWTGA SKYNLNGRSI INPGRAMSHKXDEDEFPMSGV	540
Db	481	GPCYQOQVSKTKTDNNNSNFWTGA SKYNLNGRSI INPGRAMSHKXDEDEFPMSGV	540
QY	541	MI FGKESAGASNTALDNWITDEBEIKATNPVATERFGTVA VNFQSSSTD PATGDVHAMG	600
Db	541	MI FGKESAGASNTALDNWITDEBEIKATNPVATERFGTVA VNFQSSSTD PATGDVHAMG	600
QY	601	ALPGWVWQDRVYLOGPIYMAKI PHTDGHHBSPMLMGFGELKNPFPDII LKNTPVPANPPA	660
Db	601	ALPGWVWQDRVYLOGPIYMAKI PHTDGHHBSPMLMGFGELKNPFPDII LKNTPVPANPPA	660
QY	661	EFSAIKFASPIYQYSTGVQSVIEWELQKENSKRWPEVOYJYSNVAKSANVDEPTVNNGL	720
Db	661	EFSAIKFASPIYQYSTGVQSVIEWELQKENSKRWPEVOYJYSNVAKSANVDEPTVNNGL	720
QY	721	YTEPRPIGTRYLTRPL 736	
Db	721	YTEPRPIGTRYLTRPL 736	

RESULT 3
 OS6137_gVIRU
 ID OS6137_gVIRU PRELIMINARY; PRT; 736 AA.
 AC OS6137;
 DT 01-JUN-1998, integrated into UniProtKB/TrEMBL.
 DT 01-JUN-1998, sequence version 1.
 DT 07-FEB-2006, entry version 19.
 DT Capcid protein VPI.
 OS Adeno-associated virus 6.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
 NX NCBI_TaxId=68558;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=98080418; PubMed=9420229;
 RA Rutledge B.A., Halbert C.L., Russell D.W.;
 RT "Infectious clones and vectors derived from adeno-associated virus
 (AAV) serotypes other than AAV type 2.";
 RL J. Virol. 72:309-319(1998).
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 CC EMBL; AF028704; AAB95450.1; -; Genomic_DNA.
 DR SMR; OS6137; 217-736.

DR GO: GO:0019028; C:Viral capsid; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 736 AA; 81411 MW; 311217A089C565F5 CRC64;

Query Match 99.3%; Score 3963; DB 2; Length 736;
 Best Local Similarity 99.2%; Pred. No. 2e-249;
 Matches 730; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDLSEGIKREMDLKGAPKPKANQKODDGRGLVPGYKYLDPFNGLD 60
 DB 1 MAADGYLPDWLEDLSEGIKREMDLKGAPKPKANQKODDGRGLVPGYKYLDPFNGLD 60
 QY 61 KGEVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAEPQERLQEDTSFGCNIGRAVFQ 120
 DB 61 KGEVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAEPQERLQEDTSFGCNIGRAVFQ 120
 QY 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGQOPAKKRLNFGQTDSE 180
 DB 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGQOPAKKRLNFGQTDSE 180
 QY 181 SVDPDQPLGEPPTAPPAVPTTASGGAPADNNEGADGVNAGNMHCOSTWLGDRVI 240
 DB 181 SVDPDQPLGEPPTAPPAVPTTASGGAPADNNEGADGVNAGNMHCOSTWLGDRVI 240
 QY 241 TTSTRTALPTYNHLYKQISSASTGASNDNHFGYSTPMGYDPFNRFHCHFSPRDQRL 300
 DB 241 TTSTRTALPTYNHLYKQISSASTGASNDNHFGYSTPMGYDPFNRFHCHFSPRDQRL 300
 QY 301 INNMGRPRKRLNFKLNIQYKEVTNDGVTIANNLTSTVQVFSDEYQLPYLGSAAHQ 360
 DB 301 INNMGRPRKRLNFKLNIQYKEVTNDGVTIANNLTSTVQVFSDEYQLPYLGSAAHQ 360
 QY 361 GCLPPFPADVPMIPOYGYLTLNNGSOAVGRSSFCLEYPFSQMLRTGNFTFSTFEEDVP 420
 DB 361 GCLPPFPADVPMIPOYGYLTLNNGSOAVGRSSFCLEYPFSQMLRTGNFTFSTFEEDVP 420
 QY 421 FHSSVYAHQSGLDLRLMNPFLIDQYLYLNRTONQSGSAQNKDLFSGSPAGMSVQPKMWLP 480
 DB 421 FHSSVYAHQSGLDLRLMNPFLIDQYLYLNRTONQSGSAQNKDLFSGSPAGMSVQPKMWLP 480
 QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMSCV 540
 DB 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMSCV 540
 QY 541 MIFGKESAGASNTALDNWMTDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMG 600
 DB 541 MIFGKESAGASNTALDNWMTDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMG 600
 QY 601 ALPGMWQODRDVYLQGPIMAKIPIHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPA 660
 DB 601 ALPGMWQODRDVYLQGPIMAKIPIHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPPA 660
 QY 661 EFSATKFASTITTOYSTGQSVSEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNNGL 720
 DB 661 EFSATKFASTITTOYSTGQSVSEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNNGL 720
 QY 721 YTEPRPIGRTYLRPL 736
 DB 721 YTEPRPIGRTYLRPL 736

RESULT 4
 06JUL3_9VIRU PRELIMINARY; PRT: 737 AA.
 AC 06JUL3_9VIRU
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 7.
 DE Capsid protein VP1.
 GN Name:cap1.
 OS Adeno-associated virus.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
 OC unclassified Dependovirus.
 OC NCBI_TaxID=272636;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
 RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
 RA "Clades of Adeno-associated viruses are widely disseminated in human
 tissues".
 RT J. Virol. 78:6381-6388(2004).
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonCommercial License
 CC
 DR EMBL: AY530606; AAS99291.1; -; Genomic_DNA.
 DR SMR: 06JUL3; 218-737.
 DR GO: GO:0019028; C:Viral capsid; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 737 AA; 81332 MW; 94CED0F3AFDEE492 CRC64;

Query Match 97.4%; Score 3883.5; DB 2; Length 737;
 Best Local Similarity 97.4%; Pred. No. 3e-244;
 Matches 718; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

QY 1 MAADGYLPDWLEDLSEGIKREMDLKGAPKPKANQKODDGRGLVPGYKYLDPFNGLD 60
 DB 1 MAADGYLPDWLEDLSEGIKREMDLKGAPKPKANQKODDGRGLVPGYKYLDPFNGLD 60
 QY 61 KGEVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAEPQERLQEDTSFGCNIGRAVFQ 120
 DB 61 KGEVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAEPQERLQEDTSFGCNIGRAVFQ 120
 QY 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGQOPAKKRLNFGQTDSE 179
 DB 121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGQOPAKKRLNFGQTDSE 180
 QY 181 ESVPDQPLGEPPTAPPAVPTTASGGAPADNNEGADGVNAGNMHCOSTWLGDRVI 239
 DB 181 ESVPDQPLGEPPTAPPAVPTTASGGAPADNNEGADGVNAGNMHCOSTWLGDRVI 240
 QY 240 ITSTRTALPTYNHLYKQISSASTGASNDNHFGYSTPMGYDPFNRFHCHFSPRDQRL 299
 DB 240 ITSTRTALPTYNHLYKQISSASTGASNDNHFGYSTPMGYDPFNRFHCHFSPRDQRL 300
 QY 300 LINNMGRPRKRLNFKLNIQYKEVTNDGVTIANNLTSTVQVFSDEYQLPYLGSAAHQ 359
 DB 301 LINNMGRPRKRLNFKLNIQYKEVTNDGVTIANNLTSTVQVFSDEYQLPYLGSAAHQ 360
 QY 360 GCLPPFPADVPMIPOYGYLTLNNGSOAVGRSSFCLEYPFSQMLRTGNFTFSTFEEDVP 419
 DB 361 GCLPPFPADVPMIPOYGYLTLNNGSOAVGRSSFCLEYPFSQMLRTGNFTFSTFEEDVP 420
 QY 420 PFHSSVYAHQSGLDLRLMNPFLIDQYLYLNRTONQSGSAQNKDLFSGSPAGMSVQPKMWLP 479
 DB 421 PLHSSVYAHQSGLDLRLMNPFLIDQYLYLNRTONQSGSAQNKDLFSGSPAGMSVQPKMWLP 480
 QY 480 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMSCV 539
 DB 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMSCV 540
 QY 540 VMI FGKESAGASNTALDNWMTDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAM 599
 DB 541 VMI FGKESAGASNTALDNWMTDEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAM 600
 QY 600 GALPGMWQODRDVYLQGPIMAKIPIHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPP 659
 DB 601 GALPGMWQODRDVYLQGPIMAKIPIHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPP 660
 QY 660 AFSATKFASTITTOYSTGQSVSEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVNNGL 719

```

Db      661 AERSATKFASTITQYSTGVSVIEIEMELQKENSKRNMPEVQYTSNYAKSASVDFTVDNNG 720
QY      720 LYTEPRPIGTRYLTRPL 736
      |||||
Db      721 LYTEPRPIGTRYLTRPL 737

RESULT 5
Q6JC12.9VIRU
ID Q6JC12.9VIRU PRELIMINARY; PRT; 736 AA.
AC Q6JC12;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
OC unclassified Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; AY530607; AAS9292.1; -; Genomic_DNA.
DR SMR; Q6JC12; 217-736.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81364 MW; F2D75A06662E6376 CRC64;

Query Match 97.1%; Score 3873; DB 2; Length 736;
Best Local Similarity 96.9%; Pred. No. 1.5e-243;
Matches 713; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 MAADGYLPMLENDLNSGIREMMDLKPGAKPKANQKODDGRGLVLPGYKYLGPENGDL 60
      |||||
Db 1 MAADGYLPMLEDTLSEGRQWMLKRPGRPPKPAERHKDSDRGVLVPGYKYLGPENGDL 60
      |||||
QY 61 KGEPVNAADAALAEHDKAYDQOLKAGDNPYLRVNHADAERQERLQEDTSGGNLGRAVPQ 120
      |||||
Db 61 KGEPVNAADAALAEHDKAYDQOLDSGNPVLKYNHADAERQERLQEDTSGGNLGRAVPQ 120
      |||||
QY 121 AKKRVLLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTDS 180
      |||||
Db 121 AKKRVLLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTDS 180
      |||||
QY 121 AKKRVLLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTDS 180
      |||||
Db 121 AKKRVLLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTDS 180
      |||||
QY 181 SVDPDQGLGEPAPTPAIVGPTTMASSGGAPMADNNEGADVGNASGNMHCDSWLGDVY 240
      |||||
Db 181 SVDPDQGLGEPAPTPAIVGPTTMASSGGAPMADNNEGADVGNASGNMHCDSWLGDVY 240
      |||||
QY 241 TTSTRTALPTNNHLTKOISSASTGASNNHYFGYSTPMGYFPENFFHGFSPRDNRL 300
      |||||
Db 241 TTSTRTALPTNNHLTKOISSASTGASNDNHYFGYSTPMGYFPENFFHGFSPRDNRL 300
      |||||
QY 301 INNNMGRPKRLNFKLNIQKEVTITNDGVTITANNLTSFVQVSDSEYOLPYLGSAAQ 360
      |||||
Db 301 INNNMGRPKRLNFKLNIQKEVTITNDGVTITANNLTSFVQVSDSEYOLPYLGSAAQ 360
      |||||
QY 361 GCLPPFPADVFMILPQGYLTLNNGSQAVGRSSPYCLEYFPQMLRTGNNFTFSYTFEEVP 420
      |||||
Db 361 GCLPPFPADVFMILPQGYLTLNNGSQAVGRSSPYCLEYFPQMLRTGNNFTFSYTFEEVP 420
      |||||
QY 421 FHSYIAHSQSLDRMLNPLIDQYLYLNRQTQNGSAGNKKDLLFRRGSPAGMSVQPKMWLP 480
      |||||

```

```

Db      421 FHSYIAHSQSLDRMLNPLIDQYLYLNRQTQNGSAGNKKDLLFRRGSPAGMSVQPKMWLP 480
      |||||
QY      481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESI INPGTAMASHKDEDEKFFPMGCV 540
      |||||
Db      481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESI INPGTAMASHKDEDEKFFPMGCV 540
      |||||
QY      541 MTFKESAGASNTALNNWMTDEEETKATNPVATERFGTVANFOSSSTDPAQDVHAMG 600
      |||||
Db      541 MTFKESAGASNTALNNWMTDEEETKATNPVATERFGTVANFOSSSTDPAQDVHAMG 600
      |||||
QY      601 ALPGWMOGRDYLQGPINAKIPHTDGHFHPSPIMGFGHKNPPOILIKNTFVPANPPA 660
      |||||
Db      601 ALPGWMOGRDYLQGPINAKIPHTDGHFHPSPIMGFGHKNPPOILIKNTFVPANPPA 660
      |||||
QY      661 EFSATKFASTITQYSTGVSVIEIEMELQKENSKRNMPEVQYTSNYAKSASVDFTVDNNG 720
      |||||
Db      661 EFSATKFASTITQYSTGVSVIEIEMELQKENSKRNMPEVQYTSNYAKSASVDFTVDNNG 720
      |||||
QY      721 LYTEPRPIGTRYLTRPL 736
      |||||
Db      721 LYTEPRPIGTRYLTRPL 736
      |||||

RESULT 6
Q6JC10.9VIRU
ID Q6JC10.9VIRU PRELIMINARY; PRT; 736 AA.
AC Q6JC10;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
OC unclassified Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; AY530609; AAS9294.1; -; Genomic_DNA.
DR SMR; Q6JC10; 217-736.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81388 MW; BA9463B970028DF0 CRC64;

Query Match 96.6%; Score 3853; DB 2; Length 736;
Best Local Similarity 96.6%; Pred. No. 2.9e-242;
Matches 711; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 1 MAADGYLPMLENDLNSGIREMMDLKPGAKPKANQKODDGRGLVLPGYKYLGPENGDL 60
      |||||
Db 1 MAADGYLPMLEDTLSEGRQWMLKRPGRPPKPAERHKDSDRGVLVPGYKYLGPENGDL 60
      |||||
QY 61 KGEPVNAADAALAEHDKAYDQOLKAGDNPYLRVNHADAERQERLQEDTSGGNLGRAVPQ 120
      |||||
Db 61 KGEPVNAADAALAEHDKAYDQOLDSGNPVLKYNHADAERQERLQEDTSGGNLGRAVPQ 120
      |||||
QY 121 AKKRVLLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTDS 180
      |||||
Db 121 AKKRVLLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTDS 180
      |||||
QY 121 AKKRVLLEPLGLVEGAKTAPGKKRPVQSPQEPDSSSGIGKTGOQPAKKRLNFGQTDS 180
      |||||

```



```

QY 181 SVDPQPLGEPAPPAVAGPTTMASSGGGAPMADNNEGADGVGNASGNHCDSTWLGDRYI 240
D 181 SVDPQPLGEPAPPAVAGPTTMASSGGGAPMADNNEGADGVGNASGNHCDSTWLGDRYI 240
QY 241 TTSRTMALPTYNNHLYKQISSASTGASNDNHYFGYSTPMGCFDPRFHCHFSPPDMQRL 300
D 241 TTSRTMALPTYNNHLYKQISSASTGASNDNHYFGYSTPMGCFDPRFHCHFSPPDMQRL 300
QY 301 INNMWGFPRKRLNPLFNIQYKEVTTNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQ 360
D 301 INNMWGFPRKRLNPLFNIQYKEVTTNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQ 360
QY 361 GCLPPFPADVPMIPOYGYLTNNGSAVGRSSFYCLEYFPQSOMLRTGNNFTFSYTFEEVP 420
D 361 GCLPPFPADVPMIPOYGYLTNNGSAVGRSSFYCLEYFPQSOMLRTGNNFTFSYTFEEVP 420
QY 421 FHSYSAHSQSLDRMLNPLIDQYLYLNRTQNSGSAQKDLIFSRSFAGMSVOPKXWLP 480
D 421 LHSACASQSLDRMLNPLIDQYLYLNRTQNSGSAQKDLIFSRSFAGMSVOPKXWLP 480
QY 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
D 481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 540
QY 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAMG 600
D 541 MIFKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAMG 600
QY 601 ALPGMWQODRDVYLQGPIMAKIPHTDGHFSPPLMGSGGLKNPPQILIKNTPVPANPAP 660
D 601 ALPGMWQODRDVYLQGPIMAKIPHTDGHFSPPLMGSGGLKNPPQILIKNTPVPANPAP 660
QY 661 EFSATKPAFTTQYSTGQVSEIEMELOKENSKRNPVEQYTSNYSASAVDFTVDNNGL 720
D 661 EFSATKPAFTTQYSTGQVSEIEMELOKENSKRNPVEQYTSNYSASAVDFTVDNNGL 720
QY 721 YTEPRPIGTRYLTRPL 736
D 721 YTEPRPIGTRYLTRPL 736

```

RESULT 7
056139_9VIRU PRELIMINARY; PRT; 736 AA.
AC 056139;
DT 01-JUN-1998, integrated into UniProtKB/TrEMBL.
DT 01-JUN-1998, sequence version 1.
DT 07-FEB-2006, entry version 20.
DE Capsid protein VPI.
OS Adeno-associated virus 3B.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=68742;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98080418; PubMed=9420229;
RA Rutledge E.A., Halbert C.L., Ruseell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus
J. Virol. 72:309-319 (1998)."
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial license
CC -----
DR EMBL: AF028705; AAB95452.1; -; Genomic_DNA.
DR SMR: 056139; 217-736.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005158; F:structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81907 MW; D52331AD5F0D70F CRC64;

Query March 86.0%; Score 3511; DB 2; Length 736;
Beet Local Similarity 86.8%; Pred. NO. 5.7e-220;

```

Matches 640; Conservative 38; Mismatches 57; Indels 2; Gaps 2;
QY 1 MAAGYLPDWLEDNLSGIRREMWDLKPGAPRKANQKQKODGRLVLPGYKYLGPNGLD 60
D 1 MAAGYLPDWLEDNLSGIRREMWDLKPGAPRKANQKQKODGRLVLPGYKYLGPNGLD 60
QY 61 KGEYNNADAALAEHDKAYDQOLKAGNPNYLRNHADAEFQERLEDETSFGNLCRAVFO 120
D 61 KGEYNNADAALAEHDKAYDQOLKAGNPNYLRNHADAEFQERLEDETSFGNLCRAVFO 120
QY 121 AKKVLBEPLGLVEBAGAKTAPGKKRPVEQSPQEPDSSSGIGTGOOPAKKRLNFGQTGSE 180
D 121 AKKVLBEPLGLVEBAGAKTAPGKKRPVDSQSPQEPDSSSGVGSCKOPAKKRLNFGQTGSE 180
QY 181 SVDPQPLGEPAPPAVAGPTTMASSGGGAPMADNNEGADGVGNASGNHCDSTWLGDRYI 240
D 181 SVDPQPLGEPAPPAVAGPTTMASSGGGAPMADNNEGADGVGNASGNHCDSTWLGDRYI 240
QY 241 TTSRTMALPTYNNHLYKQISSASTGASNDNHYFGYSTPMGCFDPRFHCHFSPPDMQRL 300
D 241 TTSRTMALPTYNNHLYKQISSASTGASNDNHYFGYSTPMGCFDPRFHCHFSPPDMQRL 300
QY 301 INNMWGFPRKRLNPLFNIQYKEVTTNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQ 360
D 301 INNMWGFPRKRLNPLFNIQYKEVTTNDGVTTIANNLSTVQVPSDSEYQLPYVLGSAHQ 360
QY 361 GCLPPFPADVPMIPOYGYLTNNGSAVGRSSFYCLEYFPQSOMLRTGNNFTFSYTFEEVP 420
D 361 GCLPPFPADVPMIPOYGYLTNNGSAVGRSSFYCLEYFPQSOMLRTGNNFTFSYTFEEVP 420
QY 421 FHSYSAHSQSLDRMLNPLIDQYLYLNRTQNSGSAQKDLIFSRSFAGMSVOPKXWLP 480
D 421 FHSYSAHSQSLDRMLNPLIDQYLYLNRTQNSGSAQKDLIFSRSFAGMSVOPKXWLP 480
QY 480 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 539
D 480 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGCV 539
QY 540 VMIKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAM 599
D 540 VMIKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAM 599
QY 599 NLIKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAM 659
D 599 NLIKESAGASNTALDNVMTDEBEIKATNPVATERFGTVAVNFQSSSTDPAIGDVHAM 659
QY 600 GALPGMWQODRDVYLQGPIMAKIPHTDGHFSPPLMGSGGLKNPPQILIKNTPVPANPAP 660
D 600 GALPGMWQODRDVYLQGPIMAKIPHTDGHFSPPLMGSGGLKNPPQILIKNTPVPANPAP 660
QY 660 AFSATKPAFTTQYSTGQVSEIEMELOKENSKRNPVEQYTSNYSASAVDFTVDNNG 719
D 660 AFSATKPAFTTQYSTGQVSEIEMELOKENSKRNPVEQYTSNYSASAVDFTVDNNG 719
QY 720 LYTEPRPIGTRYLTRPL 736
D 720 LYTEPRPIGTRYLTRPL 736
QY 720 YSEPRPIGTRYLTRPL 736
D 720 YSEPRPIGTRYLTRPL 736

```

RESULT 8
056311_9VIRU PRELIMINARY; PRT; 736 AA.
AC 056311;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Capsid protein.
OS Adeno-associated virus - 3.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=46350;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3H;
RX MEDLINE=96266430; PubMed=8661429; DOI=10.1006/viro.1996.0367;
RA Muramatsu S., Mizukami H., Young N.S., Brown K.E.;
RT "Nucleotide sequencing and generation of an infectious clone of adeno-associated virus 3.";

```
RL virology 221:208-217(1996).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: U48704; AAC55049.1; -; Genomic_DNA.
DR SMR: O65311; 217-736.
DR GO: GO:0019028; C:Viral capsid; IEA.
DR GO: GO:0005198; F:Structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81660 MW; APL1EF47B5C67A10 CRC64;

Query Match      87.6%; Score 3494; DB 2; Length 736;
Best Local Similarity 86.4%; Pred. No. 7.3e-219;
Matches 637; Conservative 39; Mismatches 59; Indels 2; Gaps 2;

QY 1 MAADGYLPDWLENNLSGIREMMDLKPGAPKPAKNOQKODDGRGLVLPGYKYLGPFFVGLD 60
DB 1 MAADGYLPDWLENNLSGIREMMDLKPGAPKPAKNOQKODDGRGLVLPGYKYLGPFFVGLD 60
QY 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAEOERLOEDTSFGNIGRAVFO 120
DB 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAEOERLOEDTSFGNIGRAVFO 120
QY 121 AKKRVLEPLGLVEBGAKTAPGKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGSE 180
DB 121 AKKRVLEPLGLVEBGAKTAPGKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGSE 180
QY 121 AKKRVLEPLGLVEBGAKTAPGKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGSE 180
DB 121 AKKRVLEPLGLVEBGAKTAPGKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGSE 180
QY 181 SVDDPOLGEPPTPAVGPFTTMASSGGAPMADNNEGADGVNAGSNWGHCDSTWLGDRVI 240
DB 181 SVDDPOLGEPPTPAVGPFTTMASSGGAPMADNNEGADGVNAGSNWGHCDSTWLGDRVI 240
QY 241 TTSRTMALPTYNHLYKQISASTGASNDNHFGYSTPMDYFENRFGHCFSPRDMQRL 300
DB 241 TTSRTMALPTYNHLYKQISASTGASNDNHFGYSTPMDYFENRFGHCFSPRDMQRL 300
QY 301 INNNNGRRPRKLNFLKFNIOVEKYTTNDGYTTIANNTSTVOYFSDSEYOLPYVLGSAHQ 360
DB 301 INNNNGRRPRKLNFLKFNIOVEKYTTNDGYTTIANNTSTVOYFSDSEYOLPYVLGSAHQ 360
QY 300 INNNNGRRPRKLNFLKFNIOVEKYTTNDGYTTIANNTSTVOYFSDSEYOLPYVLGSAHQ 359
DB 300 INNNNGRRPRKLNFLKFNIOVEKYTTNDGYTTIANNTSTVOYFSDSEYOLPYVLGSAHQ 359
QY 361 GCLPPFPADVFMIPQYGYLTLLNGSOAVGRSSFCLEYPSPOMLRGTGNNFTFSYTFEVP 420
DB 360 GCLPPFPADVFMIPQYGYLTLLNGSOAVGRSSFCLEYPSPOMLRGTGNNFTFSYTFEVP 419
QY 421 FHSVYAHSGSLDRIMNPLIDQYLYLNRTO-NOSGAONKDLFSRSGSPAGMSVQPRNWL 479
DB 420 FHSVYAHSGSLDRIMNPLIDQYLYLNRTO-NOSGAONKDLFSRSGSPAGMSVQPRNWL 479
QY 480 PGECYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESITNPGTAMASHKDDDEKFFPMSC 539
DB 480 PGECYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESITNPGTAMASHKDDDEKFFPMSC 539
QY 480 PGECYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESITNPGTAMASHKDDDEKFFPMSC 539
DB 480 PGECYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESITNPGTAMASHKDDDEKFFPMSC 539
QY 540 VMIFGKESAGASNTALDNWMITDEBEIKATNPVATERFGTVAVNFOSSTDPATGDVHAM 599
DB 540 VMIFGKESAGASNTALDNWMITDEBEIKATNPVATERFGTVAVNFOSSTDPATGDVHAM 599
QY 540 NLIIFGKESGTTASNAELDNWMITDEBEIKATNPVATERFGTVAVNFOSSTDPATGDVHAM 599
DB 540 NLIIFGKESGTTASNAELDNWMITDEBEIKATNPVATERFGTVAVNFOSSTDPATGDVHAM 599
QY 600 GALPGWVQDRDYLYQSPRIWAKIPHTDGHFHPSELMGGFGIKNPPQILIKNTVPVNPAP 659
DB 600 GALPGWVQDRDYLYQSPRIWAKIPHTDGHFHPSELMGGFGIKNPPQILIKNTVPVNPAP 659
QY 660 AESSATFASFTIOYSTGVSVLEIEMELOKENSGRANPEVOYTSNYAKSANVDPTDUNG 719
DB 660 AESSATFASFTIOYSTGVSVLEIEMELOKENSGRANPEVOYTSNYAKSANVDPTDUNG 719
QY 660 TTSPAPAFPIIOYSTGVSVLEIEMELOKENSGRANPEVOYTSNYAKSANVDPTDUNG 719
DB 660 TTSPAPAFPIIOYSTGVSVLEIEMELOKENSGRANPEVOYTSNYAKSANVDPTDUNG 719
QY 720 LYTEPRPIGTRYLRLPL 736
DB 720 LYTEPRPIGTRYLRLPL 736
QY 720 VISEPRPIGTRYLRLPL 736
DB 720 VISEPRPIGTRYLRLPL 736
```

```
AC Q808Y3;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE Capsid protein.
DE Name=VP1.
OS Non-human primate Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
OC unclassified Dependovirus.
OX NCBI_TaxID=226582;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22631657; Pubmed=12716974; DOI=10.1073/pnas.0937739100;
RA Gao G., Alvirra M.R., Somanathan S., Lu Y., Vandenberghe L.H.,
RA Rux J.J., Calcedo R., Samiuel J., Abbas Z., Wilson J.M.;
RT "Adeno-associated viruses undergo substantial evolution in primates
RT during natural infections."
RL Proc. Natl. Acad. Sci. U.S.A. 100:6081-6086(2003).

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AY242997; AA088183.1; -; Genomic_DNA.
DR SMR: Q808Y3; 217-736.
DR GO: GO:0019028; C:Viral capsid; IEA.
DR GO: GO:0005198; F:Structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81370 MW; 87D05047FEFAA14C CRC64;

Query Match      87.1%; Score 3473; DB 2; Length 736;
Best Local Similarity 85.6%; Pred. No. 1.7e-217;
Matches 631; Conservative 44; Mismatches 60; Indels 2; Gaps 2;

QY 1 MAADGYLPDWLENNLSGIREMMDLKPGAPKPAKNOQKODDGRGLVLPGYKYLGPFFVGLD 60
DB 1 MAADGYLPDWLENNLSGIREMMDLKPGAPKPAKNOQKODDGRGLVLPGYKYLGPFFVGLD 60
QY 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAEOERLOEDTSFGNIGRAVFO 120
DB 61 KGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAEOERLOEDTSFGNIGRAVFO 120
QY 121 AKKRVLEPLGLVEBGAKTAPGKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGSE 180
DB 121 AKKRVLEPLGLVEBGAKTAPGKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGSE 180
QY 121 AKKRVLEPLGLVEBGAKTAPGKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGSE 180
DB 121 AKKRVLEPLGLVEBGAKTAPGKRPVEQSPQEPDSSSGIGKTGOOPAKKRLNFGQTGSE 180
QY 181 SVDDPOLGEPPTPAVGPFTTMASSGGAPMADNNEGADGVNAGSNWGHCDSTWLGDRVI 240
DB 181 SVDDPOLGEPPTPAVGPFTTMASSGGAPMADNNEGADGVNAGSNWGHCDSTWLGDRVI 240
QY 241 TTSRTMALPTYNHLYKQISASTGAS-NDNHFGYSTPMDYFENRFGHCFSPRDMQRL 299
DB 241 TTSRTMALPTYNHLYKQISASTGAS-NDNHFGYSTPMDYFENRFGHCFSPRDMQRL 299
QY 241 TTSRTMALPTYNHLYKQISASTGAS-NDNHFGYSTPMDYFENRFGHCFSPRDMQRL 299
DB 241 TTSRTMALPTYNHLYKQISASTGAS-NDNHFGYSTPMDYFENRFGHCFSPRDMQRL 299
QY 241 TTSRTMALPTYNHLYKQISASTGAS-NDNHFGYSTPMDYFENRFGHCFSPRDMQRL 300
DB 241 TTSRTMALPTYNHLYKQISASTGAS-NDNHFGYSTPMDYFENRFGHCFSPRDMQRL 300
QY 300 LNNNGRRPRKLNFLKFNIOVEKYTTNDGYTTIANNTSTVOYFSDSEYOLPYVLGSAH 359
DB 301 LNNNGRRPRKLNFLKFNIOVEKYTTNDGYTTIANNTSTVOYFSDSEYOLPYVLGSAH 360
QY 301 LNNNGRRPRKLNFLKFNIOVEKYTTNDGYTTIANNTSTVOYFSDSEYOLPYVLGSAH 360
DB 301 LNNNGRRPRKLNFLKFNIOVEKYTTNDGYTTIANNTSTVOYFSDSEYOLPYVLGSAH 360
QY 360 OGCLPPFPADVFMIPQYGYLTLLNGSOAVGRSSFCLEYPSPOMLRGTGNNFTFSYTFEVP 419
DB 361 OGCLPPFPADVFMIPQYGYLTLLNGSOAVGRSSFCLEYPSPOMLRGTGNNFTFSYTFEVP 420
QY 420 PHSVYAHSGSLDRIMNPLIDQYLYLNRTO-NOSGAONKDLFSRSGSPAGMSVQPRNWL 479
DB 421 PHSVYAHSGSLDRIMNPLIDQYLYLNRTO-NOSGAONKDLFSRSGSPAGMSVQPRNWL 479
QY 480 PGECYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESITNPGTAMASHKDDDEKFFPMSC 539
DB 480 PGECYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESITNPGTAMASHKDDDEKFFPMSC 539
QY 480 PGECYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESITNPGTAMASHKDDDEKFFPMSC 539
DB 480 PGECYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESITNPGTAMASHKDDDEKFFPMSC 539
QY 540 VMIFGKESAGASNTALDNWMITDEBEIKATNPVATERFGTVAVNFOSSTDPATGDVHAM 599
DB 540 VMIFGKESAGASNTALDNWMITDEBEIKATNPVATERFGTVAVNFOSSTDPATGDVHAM 599
```



```

DE Capsid protein VP1.
OS Name-cap;
OC Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
OC Unclassified Dependovirus.
OX NCBI_Taxid=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenbergh L.H., Alvir M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues";
RL J. Virol. 78:6381-6388(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AY530561; AAS9246.1; -; Genomic_DNA.
DR SMR: Q6UC58; 216-737.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo coat.
DR Pfam: PF00740; Parvo coat; 1.
SQ SEQUENCE 737 AA; 81486 MW; EF6B85A99F0F7CB5 CRC64;

Query Match 86.6%; Score 3455.5; DB 2; Length 737;
Best Local Similarity 85.6%; Pred. No. 2.4e-216;
Matches 632; Conservative 48; Mismatches 55; Indels 3; Gaps 3;

QY 1 MAADGYLPDWLEDTLSGIREMMDLKGAPKPKANQCKODDGRGLVLPGYKYLGPFGD 60
DB 1 MAADGYLPDWLEDTLSGIREMMDLKGAPKPKANQCKODDGRGLVLPGYKYLGPFGD 60
QY 61 KGEVNAADAAALEHDAVYDQOLKAGDNPYLRYNHADAERFOERLOEDTSFGNIGRAVFQ 120
DB 61 KGEVNAADAAALEHDAVYDQOLKAGDNPYLRYNHADAERFOERLOEDTSFGNIGRAVFQ 120
QY 121 AKRRVLEPLGLVEGAKTAPKRRPVQSPQ-EPDSSSGIGTKGQQAQAKRLNFQGTGDS 179
DB 121 AKRRVLEPLGLVEGAKTAPKRRPVQSPQ-EPDSSSGIGTKGQQAQAKRLNFQGTGDS 180
QY 180 ESWDDPQPLGEPPTPAVGETTMASSGGAPMADNNEGADGVGASGMWHDSTWLDGRV 239
DB 181 ESWDDPQPLGEPPTPAVGETTMASSGGAPMADNNEGADGVGASGMWHDSTWLDGRV 240
QY 240 ITTSTRTALPTYNHLYKQISSASAGSNDNHFGYSTPMGYDPFNRPHCHFSPRDMOR 299
DB 241 ITTSTRTALPTYNHLYKQISSASAGSNDNHFGYSTPMGYDPFNRPHCHFSPRDMOR 300
QY 300 LNNMGFRPRKRLNFKLNIQVKEVTNDGVTTIANMLTSVQVFSDEYOLPYVLGSAH 359
DB 301 LNNMGFRPRKRLNFKLNIQVKEVTNDGVTTIANMLTSVQVFSDEYOLPYVLGSAH 360
QY 360 QGCLPPPADVFMIPQGYLTLNNGSQAVGRSSFYCLEYFPSSQMLRTGNFTFSYTEEV 419
DB 361 QGCLPPPADVFMIPQGYLTLNNGSQAVGRSSFYCLEYFPSSQMLRTGNFTFSYTEEV 420
QY 420 PFSSVYHSSQSLDRMLPLIDQYLYLARFOSNAGGTAKNRELQFYQGGPTTMAEQAKNW 478
DB 421 PFSSVYHSSQSLDRMLPLIDQYLYLARFOSNAGGTAKNRELQFYQGGPTTMAEQAKNW 480
QY 479 LPGEVYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTMAASHKDEDEKFPMS 538
DB 481 LPGEVYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTMAASHKDEDEKFPMS 540
QY 539 GNWIFGESAGASTALADNMTDEEIKATNPATERFGTVANVPFSSSTDPATGYVHA 598
DB 541 GNWIFGESAGASTALADNMTDEEIKATNPATERFGTVANVPFSSSTDPATGYVHA 599
QY 599 MGALPGWVMDRDVYLGPIWAKIPHTDGHFHPRLKMGFGKLPKPPQIILKNTPVANP 658
DB 600 MGALPGWVMDRDVYLGPIWAKIPHTDGHFHPRLKMGFGKLPKPPQIILKNTPVANP 659

```

```

QY 659 PAEFSATKPAFTIOYSTGVSAIEIEMLEKENSKRANPEVQYTSNKAANDPTVDN 718
DB 660 PEVFTPAKFASFIOYSTGVSAIEIEMLEKENSKRANPEIQTYSNPKQGVDAVDSQ 719
QY 719 GLYTEPRPIGRPLTRPL 736
DB 720 GVYSEPRPIGRPLTRNL 737

RESULT 14
06JC28 9VIRU PRELIMINARY; PRT; 735 AA.
AC Q6JC28;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Capsid protein VP1.
GN Name-cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
OC Unclassified Dependovirus.
OX NCBI_Taxid=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenbergh L.H., Alvir M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues";
RL J. Virol. 78:6381-6388(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AY530561; AAS9246.1; -; Genomic_DNA.
DR SMR: Q6UC28; 217-735.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR001403; Parvo coat.
DR Pfam: PF00740; Parvo coat; 1.
SQ SEQUENCE 735 AA; 81898 MW; 197066F0911FA5B9B CRC64;

Query Match 86.6%; Score 3453.5; DB 2; Length 735;
Best Local Similarity 84.5%; Pred. No. 3.2e-216;
Matches 622; Conservative 48; Mismatches 65; Indels 1; Gaps 1;

QY 1 MAADGYLPDWLEDTLSGIREMMDLKGAPKPKANQCKODDGRGLVLPGYKYLGPFGD 60
DB 1 MAADGYLPDWLEDTLSGIREMMDLKGAPKPKAERHKQSGRGLVLPGYKYLGPFGD 60
QY 61 KGEVNAADAAALEHDAVYDQOLKAGDNPYLRYNHADAERFOERLOEDTSFGNIGRAVFQ 120
DB 61 KGEVNAADAAALEHDAVYDQOLKAGDNPYLRYNHADAERFOERLOEDTSFGNIGRAVFQ 120
QY 121 AKRRVLEPLGLVEGAKTAPKRRPVQSPQ-EPDSSSGIGTKGQQAQAKRLNFQGTGDS 180
DB 121 AKRRVLEPLGLVEGAKTAPKRRPVQSPQ-EPDSSSGIGTKGQQAQAKRLNFQGTGDS 180
QY 181 ESWDDPQPLGEPPTPAVGETTMASSGGAPMADNNEGADGVGASGMWHDSTWLDGRV 240
DB 181 ESWDDPQPLGEPPTPAVGETTMASSGGAPMADNNEGADGVGASGMWHDSTWLDGRV 240
QY 241 TTSTRTALPTYNHLYKQISSASAGSNDNHFGYSTPMGYDPFNRPHCHFSPRDMOR 300
DB 241 TTSTRTALPTYNHLYKQISSASAGSNDNHFGYSTPMGYDPFNRPHCHFSPRDMOR 299
QY 301 LNNMGFRPRKRLNFKLNIQVKEVTNDGVTTIANMLTSVQVFSDEYOLPYVLGSAH 360
DB 300 LNNMGFRPRKRLNFKLNIQVKEVTNDGVTTIANMLTSVQVFSDEYOLPYVLGSAH 359
QY 361 GCLPPPADVFMIPQGYLTLNNGSQAVGRSSFYCLEYFPSSQMLRTGNFTFSYTEEV 420
DB 361 GCLPPPADVFMIPQGYLTLNNGSQAVGRSSFYCLEYFPSSQMLRTGNFTFSYTEEV 420

```

```

Db      360 GCLPPFPADVFMVPOYGYLTIANNGSOAVGSSPYCLEYFPSPQMLRTGNNFQSYTFEDV 419
Qy      421 FHSSVYAHQSGLDRLMNPILIDQYLYLNRTQNGSQAQNKDILFSRGSPPAGMSVOPKXWLP 480
Db      420 FHSSVYAHQSGLDRLMNPILIDQYLYLNRTQNGSQAQNKDILFSRGSPPAGMSVOPKXWLP 479
Qy      481 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEFFPMGSV 540
Db      480 GPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEFFPMGSV 539
Qy      541 MIRGKSAGASANTALDNVMTTDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAM 600
Db      540 LIFGKGCTANADDLNEMWITDEBEIRKTNPVATEQGYTSNNLQNSNIGPTTGTVNHG 599
Qy      601 ALPGMWQDRDVLVLOGPYIAKIPHTDGHFHPSPLMGFGGLKPNPPOILLKNTVPANPA 660
Db      600 ALPGMWQDRDVLVLOGPYIAKIPHTDGHFHPSPLMGFGGLKPNPPOILLKNTVPANPA 659
Qy      661 EFSAATFASFTQYSTQGVSEIEMELQKNSKRNPEVOYTSNVAKSANVDFTVNNG 720
Db      660 NFSSAKFASFTQYSTQGVSEIEMELQKNSKRNPEIQTSTNYSKVNVDFTVNNGV 719
Qy      721 YTEPRPIGTGYLTRPL 736
Db      720 YSEPRPIGTGYLTRNL 735

RESULT 15
Q6JC19_9VIRU PRELIMINARY; PRT; 738 AA.
ID 06JC19_9VIRU
AC 06JC19_
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DE 07-FEB-2006, entry version 7.
DE Capsid protein VP1.
GN Name:cap.
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus;
OC unclassified Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC
CC -----
DR EMBL, AY530600; AAS9285.1; -; Genomic_DNA.
DR SMR, Q6JC19; 218-738.
DR GO, GO:0019028; C:viral capsid; IEA.
DR GO, GO:0005198; F:structural molecule activity; IEA.
DR InterPro, IPR001403; Parvo_coat.
DR Pfam, PF00740; Parvo_coat; 1.
SQ SEQUENCE 738 AA; 81582 MW; 585FA6CF0769934F CRC64;

Query Match      86.6%; Score 3453; DB 2; Length 738;
Best Local Similarity 85.0%; Pred. No. 3.5e-216;
Matches 627; Conservative 48; Mismatches 61; Indels 2; Gaps 2;

```

```

Db      121 AKKRVLEPGLIVEAGTAPGKKRPVPSQ-EPDSSSGIGKTGQPAKKRIINFGQTGDS 180
Qy      180 ESVVPDQPIGEPPATPAVAGPTTMAAGGAPADNNEGADGVGNASGNWCHDSITWIGDRV 229
Db      181 ESVVPDQPIGEPPAGSGGSGTMAAGGAPADNNEGADGVGVSSSGNWHCHDSITWIGDRV 240
Qy      240 ITTSTRTMALPTYNHLYQISASGTAS-NDNHYFGYSTPMGYPFENRPHCFSPRDMQ 298
Db      241 ITTSTRTMALPTYNHLYQISNGISGSTDNDTITGYSTPMGYFENRPHCFSPRDMQ 300
Qy      299 RLINNMGRPKALNFKLNIQYKEVTNDGVTITANNLTSTVQVSEDEYOLPYVLGSA 358
Db      301 RLINNMGRPKRSLKLNIGYKEVTQNEGKTIANNLSTIQVFTDSEYOLPYVLGSA 360
Qy      359 HOGCLPPFPADVFMVPOYGYLTIANNGSOAVGSSPYCLEYFPSPQMLRTGNNFQSYTFEE 418
Db      361 HOGCLPPFPADVFMVPOYGYLTIANNGSOAVGSSPYCLEYFPSPQMLRTGNNFQSYTFED 420
Qy      419 VPFHSSVYAHQSGLDRLMNPILIDQYLYLNRTQNGSQAQNKDILFSRGSPPAGMSVOPKXW 478
Db      421 VPFHSSVYAHQSGLDRLMNPILIDQYLYLNRTQNGSQAQNKDILFSRGSPPAGMSVOPKXW 480
Qy      479 LPGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEFFPM 538
Db      481 LPGPCYRQORVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEFFPM 540
Qy      539 GVMIPEGKESAGASNTALDNVMTTDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDVHA 598
Db      541 GVMIPEGKESAGASNTALDNVMTTDEBEIKATNPVATERFGTVAVNFQSSSTDPATGDVHA 600
Qy      599 MGALPGMWQDRDVLVLOGPYIAKIPHTDGHFHPSPLMGFGGLKPNPPOILLKNTVPANP 658
Db      601 MGALPGMWQDRDVLVLOGPYIAKIPHTDGHFHPSPLMGFGGLKPNPPOILLKNTVPANP 660
Qy      659 PAEFSATKFAFTQYSTQGVSEIEMELQKNSKRNPEVOYTSNVAKSANVDFTVNNG 718
Db      661 PTFESQAKLASFTQYSTQGVSEIEMELQKNSKRNPEIQTSTNYSKSTNVDFAVNTE 720
Qy      719 GLYTEPRPIGTGYLTRPL 736
Db      721 GYTEPRPIGTGYLTRNL 738

```

Search completed: June 12, 2006, 12:52:24
 Job time : 308 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2006, 12:52:39 ; Search time 52 Seconds
(without alignments)
1238.895 Million cell updates/sec

Title: US-10-696-261-13

Perfect score: 3989
Sequence: 1 MAADGYLPMLENDLSEGR.....NGLYTEPRPIGRYLRPL 736

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /EMC_Celerra_SIDS3/prodata/2/1aa/5.COMB.pep:*
2: /EMC_Celerra_SIDS3/prodata/2/1aa/6.COMB.pep:*
3: /EMC_Celerra_SIDS3/prodata/2/1aa/7.COMB.pep:*
4: /EMC_Celerra_SIDS3/prodata/2/1aa/H.COMB.pep:*
5: /EMC_Celerra_SIDS3/prodata/2/1aa/PCTUS.COMB.pep:*
6: /EMC_Celerra_SIDS3/prodata/2/1aa/RE.COMB.pep:*
7: /EMC_Celerra_SIDS3/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3989	100.0	736	2	US-09-807-802A-3 Sequence 3, Appl1
2	3989	100.0	736	2	US-09-807-802A-13 Sequence 13, Appl1
3	3402.5	85.3	735	2	US-09-321-589-1 Sequence 1, Appl1
4	3402.5	85.3	735	2	US-10-293-478-1 Sequence 1, Appl1
5	3402.5	85.3	735	2	US-10-038-972A-13 Sequence 13, Appl1
6	3402.5	85.3	735	2	US-09-717-789C-24 Sequence 24, Appl1
7	3251	81.5	599	2	US-09-807-802A-15 Sequence 15, Appl1
8	2906	72.9	534	2	US-09-807-802A-17 Sequence 17, Appl1
9	2759.5	69.2	538	2	US-10-038-972A-14 Sequence 14, Appl1
10	2486.5	62.2	533	2	US-09-532-594B-4 Sequence 4, Appl1
11	2481.5	62.2	533	2	US-10-038-972A-15 Sequence 15, Appl1
12	2220	55.7	724	2	US-09-533-427-4 Sequence 4, Appl1
13	2220	55.7	724	2	US-09-717-789C-4 Sequence 4, Appl1
14	1830.5	45.9	588	2	US-09-532-594B-16 Sequence 16, Appl1
15	1700.5	42.6	588	2	US-09-533-427-5 Sequence 5, Appl1
16	1700.5	42.6	588	2	US-09-717-789C-5 Sequence 5, Appl1
17	1690.5	42.4	544	2	US-09-532-594B-18 Sequence 18, Appl1
18	1665	41.7	532	2	US-09-533-427-6 Sequence 6, Appl1
19	1665	41.7	532	2	US-09-717-789C-6 Sequence 6, Appl1
20	1410	35.3	756	2	US-09-438-268-4 Sequence 4, Appl1
21	599.5	15.0	781	2	US-10-187-253E-27 Sequence 27, Appl1
22	590.5	14.8	781	2	US-10-187-253E-33 Sequence 33, Appl1
23	486	12.2	554	2	US-10-187-253E-29 Sequence 29, Appl1
24	486	12.2	554	2	US-10-187-253E-35 Sequence 35, Appl1
25	479.5	12.0	543	2	US-08-856-841-22 Sequence 22, Appl1
26	430.5	10.8	500	2	US-08-856-841-16 Sequence 16, Appl1

27	430.5	10.8	501	2	US-08-856-841-18 Sequence 18, Appl1
28	429.5	10.8	486	2	US-08-856-841-19 Sequence 19, Appl1
29	326	8.2	415	2	US-08-856-841-20 Sequence 20, Appl1
30	318	8.0	395	2	US-08-856-841-13 Sequence 13, Appl1
31	310	7.8	264	2	US-08-856-841-13 Sequence 14, Appl1
32	303	7.6	398	2	US-08-856-841-21 Sequence 21, Appl1
33	284	7.1	387	2	US-08-856-841-17 Sequence 17, Appl1
34	216.5	5.4	579	2	5223424-13 Patent No. 5223424
35	192	4.8	584	2	US-09-022-949-2 Sequence 2, Appl1
36	142	3.6	210	2	US-08-856-841-9 Sequence 9, Appl1
37	142	3.6	227	2	US-08-856-841-15 Sequence 15, Appl1
38	142	3.6	250	2	US-08-856-841-12 Sequence 12, Appl1
39	132.5	3.3	489	2	US-10-376-397B-4 Sequence 4, Appl1
40	132.5	3.3	489	2	US-10-889-526-4 Sequence 4, Appl1
41	132.5	3.3	489	3	US-10-889-527-4 Sequence 4, Appl1
42	127	3.2	1287	2	US-10-094-749-2627 Sequence 2627, Ap
43	124	3.1	3060	1	US-08-487-826B-14 Sequence 14, Appl1
44	122.5	3.1	1637	2	US-10-172-502-14 Sequence 14, Appl1
45	121	3.0	2736	2	US-09-252-991A-30227 Sequence 30227, A

ALIGNMENTS

```
RESULT 1
US-09-807-802A-3
; Sequence 3, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807, 802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR FILING DATE: 1998-11-05
; PRIOR FILING DATE: 1999-11-02
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-3

Query Match      100.0%; Score 3989; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGYLPMLENDLSEGRWMDLKPGRKPKRANQKODGGLVPGYKYLGPFGLD 60
1 MAADGYLPMLENDLSEGRWMDLKPGRKPKRANQKODGGLVPGYKYLGPFGLD 60
DB KGEFVNAAADAALEHDAVYQOLKAGDNPYLRYNHADAEFOERLOEDTSGCNGRAVFQ 120
61 KGEFVNAAADAALEHDAVYQOLKAGDNPYLRYNHADAEFOERLOEDTSGCNGRAVFQ 120
QY AKKRVLEPLGLVEBGAKTAPGKRPVQOSPOEPDSSGIGKTGOQPAKKRLNFGQTDSE 180
121 AKKRVLEPLGLVEBGAKTAPGKRPVQOSPOEPDSSGIGKTGOQPAKKRLNFGQTDSE 180
DB AKKRVLEPLGLVEBGAKTAPGKRPVQOSPOEPDSSGIGKTGOQPAKKRLNFGQTDSE 180
121 AKKRVLEPLGLVEBGAKTAPGKRPVQOSPOEPDSSGIGKTGOQPAKKRLNFGQTDSE 180
QY SVDPDPOPLGEPPTPAVGGTMASSGGGAPMAUNBEGADGVNAGNMGWHDSTLGRV 240
181 SVDPDPOPLGEPPTPAVGGTMASSGGGAPMAUNBEGADGVNAGNMGWHDSTLGRV 240
DB SVDPDPOPLGEPPTPAVGGTMASSGGGAPMAUNBEGADGVNAGNMGWHDSTLGRV 240
181 SVDPDPOPLGEPPTPAVGGTMASSGGGAPMAUNBEGADGVNAGNMGWHDSTLGRV 240
QY TTSRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPWFDPFNRFCHFSFRDQRL 300
241 TTSRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPWFDPFNRFCHFSFRDQRL 300
DB TTSRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPWFDPFNRFCHFSFRDQRL 300
241 TTSRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPWFDPFNRFCHFSFRDQRL 300
QY INNMGFRPKRLNFKLFNIOQKEVTNDGVTTIANNLITVQVFSDEYOLPYVLGSAHQ 360
301 INNMGFRPKRLNFKLFNIOQKEVTNDGVTTIANNLITVQVFSDEYOLPYVLGSAHQ 360
```



```
Db      301 INNMGRRPRKLNLFKLFNIQKEVTINDGVTIANNLSTVOQFSOSEIQLPYVLASAHQ 360
Qy      361 GCLPPPADVFMIPOGYLTLNNGSQAVGRSSFCLEYEPPSQMLRTGNFTFSYTEBEVP 420
Db      361 GCLPPPADVFMIPOGYLTLNNGSQAVGRSSFCLEYEPPSQMLRTGNFTFSYTEBEVP 420
Qy      421 FHSSVYHSOSLDRLAMPPLIDQYLYLNRTONOGSAONKDLFSRGS PAGMSVQPKWMLP 480
Db      421 FHSSVYHSOSLDRLAMPPLIDQYLYLNRTONOGSAONKDLFSRGS PAGMSVQPKWMLP 480
Qy      481 GPCYRQORVSKTIDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSV 540
Db      481 GPCYRQORVSKTIDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSV 540
Qy      541 MIFGKSAGASNTALDNVMTDEEBEIKATNPVATERFGTVAANFQSSSTD PATGDVHAMG 600
Db      541 MIFGKSAGASNTALDNVMTDEEBEIKATNPVATERFGTVAANFQSSSTD PATGDVHAMG 600
Qy      601 ALPGMWQDRDVLQGPIMAKIPIHTDGHFHPSPLMGFGGLKNPPOLLIKNTVPANPPA 660
Db      601 ALPGMWQDRDVLQGPIMAKIPIHTDGHFHPSPLMGFGGLKNPPOLLIKNTVPANPPA 660
Qy      661 EFSATKPFASFTIOYSTGOVSVEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 720
Db      661 EFSATKPFASFTIOYSTGOVSVEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 720
Qy      721 YTEPRPIGTRYLTRPL 736
Db      721 YTEPRPIGTRYLTRPL 736

RESULT 2
US-09-807-802A-13
; Sequence 13, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vector and Host Cells Containing Same
; FILE REFERENCE: GNPV.031USA
; CURRENT APPLICATION NUMBER: US/09/807, 802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 13
; LENGTH: 736
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-13

Query Match      100.0%; Score 3989; DB 2; Length 736;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAADGYLPDWLENTLSEGIREWMDLKGAPKPKANOOKODGSGVLVPGYKYLGPFGNGLD 60
Db      1 MAADGYLPDWLENTLSEGIREWMDLKGAPKPKANOOKODGSGVLVPGYKYLGPFGNGLD 60
Qy      61 KGPBVNAADAALAEHDKAYDQOLKAGDNPLYRYNHADAFOERLQEDTSGFNLGRAVFO 120
Db      61 KGPBVNAADAALAEHDKAYDQOLKAGDNPLYRYNHADAFOERLQEDTSGFNLGRAVFO 120
Qy      121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGOOPAKKALNFQOTDSE 180
Db      121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGOOPAKKALNFQOTDSE 180
Qy      181 SVDPQPLGPRPATPAVGPPTTASGGGAPMADNNEGADGVGNASGNWHODSTWLGDRVI 240
```

```
Db      181 SVDPQPLGPRPATPAVGPPTTASGGGAPMADNNEGADGVGNASGNWHODSTWLGDRVI 240
Qy      241 TTSRTWALPTYNNHLKYOISSASTGASNDNHFGYSTWGYFDNRFCHSPRDMOQL 300
Db      241 TTSRTWALPTYNNHLKYOISSASTGASNDNHFGYSTWGYFDNRFCHSPRDMOQL 300
Qy      301 INNMGRRPRKLNLFKLFNIQKEVTINDGVTIANNLSTVOQFSOSEIQLPYVLASAHQ 360
Db      301 INNMGRRPRKLNLFKLFNIQKEVTINDGVTIANNLSTVOQFSOSEIQLPYVLASAHQ 360
Qy      361 GCLPPPADVFMIPOGYLTLNNGSQAVGRSSFCLEYEPPSQMLRTGNFTFSYTEBEVP 420
Db      361 GCLPPPADVFMIPOGYLTLNNGSQAVGRSSFCLEYEPPSQMLRTGNFTFSYTEBEVP 420
Qy      421 FHSSVYHSOSLDRLAMPPLIDQYLYLNRTONOGSAONKDLFSRGS PAGMSVQPKWMLP 480
Db      421 FHSSVYHSOSLDRLAMPPLIDQYLYLNRTONOGSAONKDLFSRGS PAGMSVQPKWMLP 480
Qy      481 GPCYRQORVSKTIDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSV 540
Db      481 GPCYRQORVSKTIDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDEDEKFFPMGSV 540
Qy      541 MIFGKSAGASNTALDNVMTDEEBEIKATNPVATERFGTVAANFQSSSTD PATGDVHAMG 600
Db      541 MIFGKSAGASNTALDNVMTDEEBEIKATNPVATERFGTVAANFQSSSTD PATGDVHAMG 600
Qy      601 ALPGMWQDRDVLQGPIMAKIPIHTDGHFHPSPLMGFGGLKNPPOLLIKNTVPANPPA 660
Db      601 ALPGMWQDRDVLQGPIMAKIPIHTDGHFHPSPLMGFGGLKNPPOLLIKNTVPANPPA 660
Qy      661 EFSATKPFASFTIOYSTGOVSVEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 720
Db      661 EFSATKPFASFTIOYSTGOVSVEIEMELQKENSKRANPEVOYTSNYAKSANVDFTVDNGL 720
Qy      721 YTEPRPIGTRYLTRPL 736
Db      721 YTEPRPIGTRYLTRPL 736

RESULT 3
US-09-321-589-1
; Sequence 1, Application US/09321589
; Patent No. 6498244
; GENERAL INFORMATION:
; APPLICANT: PATEL, SATIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADEMO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/09/321,589
; CURRENT FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-09-321-589-1

Query Match      85.3%; Score 3402.5; DB 2; Length 735;
Best Local Similarity 83.3%; Pred. No. 7,9e-290;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;

Qy      1 MAADGYLPDWLENTLSEGIREWMDLKGAPKPKANOOKODGSGVLVPGYKYLGPFGNGLD 60
Db      1 MAADGYLPDWLENTLSEGIREWMDLKGAPKPKANOOKODGSGVLVPGYKYLGPFGNGLD 60
Qy      61 KGPBVNAADAALAEHDKAYDQOLKAGDNPLYRYNHADAFOERLQEDTSGFNLGRAVFO 120
Db      61 KGPBVNAADAALAEHDKAYDQOLKAGDNPLYRYNHADAFOERLQEDTSGFNLGRAVFO 120
Qy      121 AKKRVLEPLGLVEGAKTAPGKKRPVEQSPQEPDSSSGIGTKGOOPAKKALNFQOTDSE 180
```

```
Dh 121 AKKRVLEPLGVEGATPAAGKRPVEHSVEPDPSSSGTGAGGQAPARKLNFGQTDAD 180
Qy 181 SVDPQPLGEPATPAAVGPTTMASSGGAPWADNNEGADVGNAAGWCHDSTWLGDRVI 240
Dh 181 SVDPQPLGEPATPAAVGPTTMASSGGAPWADNNEGADVGNAAGWCHDSTWLGDRVI 240
Qy 241 TTSRTTALPTYNHLYKQISSASTGASNDNHFGYSTPMGYFDPNRFCHFSRDMQRL 300
Dh 241 TTSRTTALPTYNHLYKQISSASTGASNDNHFGYSTPMGYFDPNRFCHFSRDMQRL 299
Qy 301 INNMGRPRKLNFKLNIQVEYTDGVTIANNLSTVQVPSDEYOLPYLGSAGHQ 360
Dh 300 INNMGRPRKLNFKLNIQVEYTDGVTIANNLSTVQVPSDEYOLPYLGSAGHQ 359
Qy 361 GCLPPFADVFMIPOGYLTLNNGSAVGRSSFYCLEYFPSQMLRTGNFTFSYTFEEVP 420
Dh 360 GCLPPFADVFMIPOGYLTLNNGSAVGRSSFYCLEYFPSQMLRTGNFTFSYTFEEVP 419
Qy 421 FHSYASQSILDRMLNPLIDQYLYLNRTONQSSAQNKDILFBSGSPAGMSVQPKMWLP 480
Dh 420 FHSYASQSILDRMLNPLIDQYLYLNRTONQSSAQNKDILFBSGSPAGMSVQPKMWLP 479
Qy 481 GPCYRQORVSKTNDNNSNFTWASKYNLNGRESIINPCTAMASHKDEDEKFFPMQSV 540
Dh 480 GPCYRQORVSKTNDNNSNFTWASKYNLNGRESIINPCTAMASHKDEDEKFFPMQSV 539
Qy 541 MIFKESAGASNTALDNVMTDEEBEIKATNPVATEREGYAVNFQSSSTPATGDVYAMG 600
Dh 540 LIFGQSEKTNVDIEKVMITDEEBEIRTNPVATEQVGSVSTNQRGNRAATADVNTQG 599
Qy 601 ALPGMWQDDVYIQQGIYAKIPHTDGHFHPSPLMGFGKLNPPQILIKXTPVPANPRA 660
Dh 600 VLPGMWQDDVYIQQGIYAKIPHTDGHFHPSPLMGFGKLNPPQILIKXTPVPANPST 659
Qy 661 EFSATKPSFTTOYSTGOVSVEIEMELQKENSKRANPEVOYTSYAKSANVDFTVDNGL 720
Dh 660 TFSAAKPSFTTOYSTGOVSVEIEMELQKENSKRANPEIQTSTYNKSNVAVDFTVDTNGV 719
Qy 721 YTEPRPIGTRYLTRPL 736
Dh 720 YSEBRPIGTRYLTRNL 735

RESULT 4
US-10-293-478-1
; Sequence 1, Application US/10293478
; Patent No. 6733757
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALIL D.
; APPLICANT: MCARTHUR, JAMES G.
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/10/293,478
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: US/09/321,589
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-10-293-478-1

Query Match 85.3%; Score 3402.5; DB 2; Length 735;
Best Local Similarity 83.3%; Pred. No. 7.9e-290;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;
```

```
Dh 61 KGEFVNADAAALEHDAKYDQQLKAGDNPIYLRYNHADAERLQEDTSGGNIQRAVQ 120
Qy 121 AKKRVLEPLGVEGATPAAGKRPVEHSVEPDPSSSGTGAGGQAPARKLNFGQTDSE 180
Dh 121 AKKRVLEPLGVEGATPAAGKRPVEHSVEPDPSSSGTGAGGQAPARKLNFGQTDSE 180
Qy 181 SVDPQPLGEPATPAAVGPTTMASSGGAPWADNNEGADVGNAAGWCHDSTWLGDRVI 240
Dh 181 SVDPQPLGEPATPAAVGPTTMASSGGAPWADNNEGADVGNAAGWCHDSTWLGDRVI 240
Qy 241 TTSRTTALPTYNHLYKQISSASTGASNDNHFGYSTPMGYFDPNRFCHFSRDMQRL 300
Dh 241 TTSRTTALPTYNHLYKQISSASTGASNDNHFGYSTPMGYFDPNRFCHFSRDMQRL 299
Qy 301 INNMGRPRKLNFKLNIQVEYTDGVTIANNLSTVQVPSDEYOLPYLGSAGHQ 360
Dh 300 INNMGRPRKLNFKLNIQVEYTDGVTIANNLSTVQVPSDEYOLPYLGSAGHQ 359
Qy 361 GCLPPFADVFMIPOGYLTLNNGSAVGRSSFYCLEYFPSQMLRTGNFTFSYTFEEVP 420
Dh 360 GCLPPFADVFMIPOGYLTLNNGSAVGRSSFYCLEYFPSQMLRTGNFTFSYTFEEVP 419
Qy 421 FHSYASQSILDRMLNPLIDQYLYLNRTONQSSAQNKDILFBSGSPAGMSVQPKMWLP 480
Dh 420 FHSYASQSILDRMLNPLIDQYLYLNRTONQSSAQNKDILFBSGSPAGMSVQPKMWLP 479
Qy 481 GPCYRQORVSKTNDNNSNFTWASKYNLNGRESIINPCTAMASHKDEDEKFFPMQSV 540
Dh 480 GPCYRQORVSKTNDNNSNFTWASKYNLNGRESIINPCTAMASHKDEDEKFFPMQSV 539
Qy 541 MIFKESAGASNTALDNVMTDEEBEIKATNPVATEREGYAVNFQSSSTPATGDVYAMG 600
Dh 540 LIFGQSEKTNVDIEKVMITDEEBEIRTNPVATEQVGSVSTNQRGNRAATADVNTQG 599
Qy 601 ALPGMWQDDVYIQQGIYAKIPHTDGHFHPSPLMGFGKLNPPQILIKXTPVPANPRA 660
Dh 600 VLPGMWQDDVYIQQGIYAKIPHTDGHFHPSPLMGFGKLNPPQILIKXTPVPANPST 659
Qy 661 EFSATKPSFTTOYSTGOVSVEIEMELQKENSKRANPEVOYTSYAKSANVDFTVDNGL 720
Dh 660 TFSAAKPSFTTOYSTGOVSVEIEMELQKENSKRANPEIQTSTYNKSNVAVDFTVDTNGV 719
Qy 721 YTEPRPIGTRYLTRPL 736
Dh 720 YSEBRPIGTRYLTRNL 735

RESULT 5
US-10-038-972A-13
; Sequence 13, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; FILE REFERENCE: 28335/36960US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VP1 capsid protein
US-10-038-972A-13

Query Match 85.3%; Score 3402.5; DB 2; Length 735;
Best Local Similarity 83.3%; Pred. No. 7.9e-290;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;
```

```

Db 1 MAA05YLP0WMBEDT1SEGI0R0WMLK6P PP6P6A6RH0X0DS6GLV6PG6K6Y6G6P6N6GLD 60
Qy 61 KGE6P6N6A6DA6AL6EH6KA6Y6D006LK6AD6N6Y6L6R6Y6N6DA6E6F0E6RL6Q6ED6TS6FG6N6I6G6RA6V6Q 120
Db 61 KGE6P6N6A6DA6AL6EH6KA6Y6D6R06L6D6S6D6N6Y6L6K6N6H6A6E6F0E6RL6K6D6TS6I6G6N6L6G6RA6V6Q 120
Qy 121 AKK6R6V6LE6PL6GL6VE6GA6K6T6A6P6KK6R6P6VE6S6P6Q6E6P6D6S6S6I6G6K6T6Q6O6P6AK6KR6L6N6F6Q6T6D6S6 180
Db 121 AKK6R6V6LE6PL6GL6VE6BV6K6T6A6P6KK6R6P6E6H6S6P6VE6D6S6S6G6K6G6K6Q6O6P6A6R6K6L6N6F6Q6T6D6D6 180
Qy 181 SVP6D6P6Q6P6I6G6E6P6A6T6A6A6V6P6T6M6A6S6G6G6A6P6M6D6N6E6G6C6V6G6N6A6S6N6W6H6D6S6T6W6L6D6R6V6I 240
Db 181 SVP6D6P6Q6P6I6G6O6P6A6A6S6G6L6G6T6N6T6M6A6T6S6G6A6P6M6D6N6N6G6A6G6V6N6S6S6G6M6H6D6S6T6W6M6G6R6V6I 240
Qy 241 TTS6T6R6T6M6L6P6T6Y6N6N6H6L6Y6K06I6S6A6S6T6A6S6A6D6N6H6F6G6S6T6P6G6Y6D6F6N6R6F6H6F6S6P6R6D6M6O6R6L 300
Db 241 TTS6T6R6T6M6L6P6T6Y6N6N6H6L6Y6K06I6S6S06S6A6S6D6N6H6F6G6I6S6T6P6G6Y6D6F6N6R6F6H6F6S6P6R6D6M6O6R6L 300
Qy 301 INN6N6G6F6R6P6K6R6L6N6F6K6L6E6N6I6Q6K6E6V6T6T6N6D6G6V6T6T6I6A6N6L6T6S6T6V6Q6F6D6S6E6Y6L6P6V6L6S6A6H6Q 360
Db 301 INN6N6G6F6R6P6K6R6L6N6F6K6L6E6N6I6Q6K6E6V6T6N6D6G6T6T6I6A6N6L6T6S6T6V6Q6F6T6D6S6E6Y6L6P6V6L6S6A6H6Q 360
Qy 361 GCL6P6P6P6A6D6V6M6I6P6O6Y6G6Y6L6T6N6N6G6A6V6R6S6S6F6Y6C6L6E6P6S6Q6L6R6T6G6N6R6T6E6S6Y6T6E6E6V6P 420
Db 361 GCL6P6P6P6A6D6V6M6I6P6O6Y6G6Y6L6T6N6N6G6A6V6R6S6S6F6Y6C6L6E6P6S6Q6L6R6T6G6N6R6T6E6S6Y6T6E6E6V6P 420
Qy 421 FHS6S6T6A6H6S6O6S6L6R6L6N6A6P6L6I6D6Y6L6Y6Y6N6T6R6O6N6G6S6A6Q6N6K6D6L6F6S6R6S6P6A6G6S6V6Q6P6K6M6L6P 480
Db 421 FHS6S6T6A6H6S6O6S6L6R6L6N6A6P6L6I6D6Y6L6Y6Y6N6T6R6O6N6G6S6A6Q6S6I6R6O6S6R6M6L6P 480
Qy 481 GPC6Y6R6Q6R6V6S6K6T6K6T6D6N6N6S6N6F6T6W6T6G6A6K6Y6N6L6N6G6R6E6I6N6E6T6A6M6A6H6K6D6E6K6F6P6M6S6V 540
Db 481 GPC6Y6R6Q6R6V6S6K6T6S6A6D6N6N6S6E6Y6S6T6G6A6T6K6Y6H6L6N6G6R6D6L6V6N6P6G6A6M6A6H6K6D6E6K6F6P6Q6S6V 540
Qy 541 MIF6G6E6S6A6C6A6S6T6A6L6D6N6W6I6T6D6E6E6I6K6A6N6P6V6A6T6E6F6Y6A6V6N6F6O6S6S6T6D6P6A6G6D6V6A6M6G 6000
Db 541 MIF6G6E6S6A6C6A6S6T6A6L6D6N6W6I6T6D6E6E6I6K6A6N6P6V6A6T6E6F6Y6A6V6N6F6O6S6S6T6D6P6A6G6D6V6A6M6G 6000
Qy 601 ALP6G6W6O6D6R6D6Y6L6O6P6I6M6A6K6I6P6T6G6H6F6H6P6S6P6L6M6G6F6G6L6K6N6P6R6P6O6L6I6N6T6P6V6A6N6P6A 6666
Db 601 VLP6G6W6O6D6R6D6Y6L6O6P6I6M6A6K6I6P6T6G6H6F6H6P6S6P6L6M6G6F6G6L6K6N6P6R6P6O6L6I6N6T6P6V6A6N6P6S6T 6655
Qy 661 EFS6A6T6K6F6A6S6F6I6O6Y6S6T6G6V6S6E6I6E6W6L6O6K6E6N6S6K6R6N6P6E6V6O6Y6T6S6N6Y6A6S6A6N6V6D6T6V6N6G6L 720
Db 660 TFS6A6K6F6A6S6F6I6O6Y6S6T6G6V6S6E6I6E6W6L6O6K6E6N6S6K6R6N6P6E6I6G6T6S6N6Y6A6S6A6N6V6D6T6V6T6N6G6V 719
Qy 721 YTE6P6P6I6G6T6R6Y6L6TR6PL 736
Db 720 YSE6P6P6I6G6T6R6Y6L6TR6NL 735

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6984517e =
; OTHER INFORMATION: synthetic construct
US-09-717-789C-24

Query Match      85.3%; Score 3402.5; DB 2; Length 735;
Best Local Similarity 83.3%; Pred. No. 7.9e-290;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;

      1  MAADGVLPMLEEDNLSEGLREEMWDLKPGAPKPKYAKOQKDDGRGLVPGYKYLGPENGID 60
      1  MAADGVLPDLLEDTLSGIRQWKLKPGPPRPKARHKKDDSGVLPGYKYLGPENGID 60
Db      1  MAADGVLPDLLEDTLSGIRQWKLKPGPPRPKARHKKDDSGVLPGYKYLGPENGID 60

      61  KGEVNVNADAAALEHDKAYDQOLKAGDNPYLRYNHADAEOERLQEDTSPGNGLGRAVQ 120
      61  KGEVNVNADAAALEHDKAYDRLDSDGNPFLKYNHADAEOERLQEDTSPGNGLGRAVQ 120
Db      61  KGEVNVNADAAALEHDKAYDRLDSDGNPFLKYNHADAEOERLQEDTSPGNGLGRAVQ 120

      121  AKKRVLEPLGLVEGATAPGKKRPVEQSGOEQDSSSGIGKQGOQAKRLNPGQGDSE 180
      121  AKKRVLEPLGLVEEVPVTAQKKRPVEHSEVPEQDSSSGIGKQGOQAKRLNPGQGDSD 180
Db      121  AKKRVLEPLGLVEEVPVTAQKKRPVEHSEVPEQDSSSGIGKQGOQAKRLNPGQGDSD 180

      181  SVDPDQELGEPPTPAAVGPTTMASSGGGAPMADNNGAGGVGASNGNHCDSTWLGDVY 240
      181  SVDPDQELGEPPTPAVSGTGNTMATSGAPMADNNGAGGVGASNGNHCDSTWMDRY 240
Db      181  SVDPDQELGEPPTPAVSGTGNTMATSGAPMADNNGAGGVGASNGNHCDSTWMDRY 240

      241  TTSTRVWALPTVNNHLKYKQISSASTGASNDNHFGYSTWGYFDENRFCHFSPRDMQRL 300
      241  TTSTRVWALPTVNNHLKYKQISSQS-GASNDNHFGYSTWGYFDENRFCHFSPRDMQRL 299
Db      241  TTSTRVWALPTVNNHLKYKQISSQS-GASNDNHFGYSTWGYFDENRFCHFSPRDMQRL 299

      301  INNNWGRPRKLNPKFLNIOVKERTNDGTTIANLSTVQVPSQSEYOLPVLGSAHQ 360
      301  INNNWGRPRKLNPKFLNIOVKERTNDGTTIANLSTVQVFTDSEYOLPVLGSAHQ 359
Db      301  INNNWGRPRKLNPKFLNIOVKERTNDGTTIANLSTVQVFTDSEYOLPVLGSAHQ 359

      361  GCLPPPADVPMIPQVGYLTNNGSAQVGRSFCLEYPFSQMLRTGNNTFESYEEVP 420
      361  GCLPPPADVPMIPQVGYLTNNGSAQVGRSFCLEYPFSQMLRTGNNTFTSYTEDVP 419
Db      361  GCLPPPADVPMIPQVGYLTNNGSAQVGRSFCLEYPFSQMLRTGNNTFTSYTEDVP 419

      421  FHSYSYASQSLDRMLNPLDQVLYLNRTOGSAQNDKLLFSRSGPAGMSVQPNWLP 480
      421  FHSYSYASQSLDRMLNPLDQVLYLNRTRTPGTTQSRLLQSSAGASDIRQSHNWL 479
Db      421  FHSYSYASQSLDRMLNPLDQVLYLNRTRTPGTTQSRLLQSSAGASDIRQSHNWL 479

      481  GPCYRQORVSKTLDNNNSNFTTGASKYLNLRRESIINPGTMAHKKDDEDFPFMSGV 540
      481  GPCYRQORVSKTLDNNNSSEYWTGATKYLNRDRLVNGPMAHKKDDEKFFPQSV 539
Db      481  GPCYRQORVSKTLDNNNSSEYWTGATKYLNRDRLVNGPMAHKKDDEKFFPQSV 539

      541  MIPKESAGSNLTADVMITDEEELKATNPVATESPGTAAVAFQSSSTDPATGDVHANG 600
      541  LIFKQSGSEKINDIEKMITDEEELKATNPVATESPGTAAVAFQSSSTDPATGDVHANG 599
Db      541  LIFKQSGSEKINDIEKMITDEEELKATNPVATESPGTAAVAFQSSSTDPATGDVHANG 599

      601  ALPGMWQDDVYLGCPIMAKIPHTDGHFHPSLMGFGILKNPPOILLKATPVNPANPA 660
      601  VLPWMQDDVYLGCPIMAKIPHTDGHFHPSLMGFGILKNPPOILLKATPVNPANPST 659
Db      601  VLPWMQDDVYLGCPIMAKIPHTDGHFHPSLMGFGILKNPPOILLKATPVNPANPST 659

      661  EFSATKTSASFTTQYSTQGVSEIEMELQKENSGRKNPEVQYTSNYAKSANDFTVNNGL 720
      661  TFSAKTSASFTTQYSTQGVSEIEMELQKENSGRKNPEIQTYSNYSKSVNVDFVTDNGV 719
Db      661  TFSAKTSASFTTQYSTQGVSEIEMELQKENSGRKNPEIQTYSNYSKSVNVDFVTDNGV 719

      721  YTEPRPIGTRYLRPL 736
      721  YSEPRPIGTRYLRPL 735
Db      721  YSEPRPIGTRYLRPL 735

RESULT 7
US-09-807-802A-15
; Sequence 15, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA

```

```

; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 599
; TYPE: PR
; ORGANISM: AAV-1
US-09-807-802A-15
```

```

Query Match      81.5%; Score 3251; DB 2; Length 599;
Best Local Similarity 100.0%; Pred. No. 1,2e-276;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      138  TAPGKRPVEQSPOEDSSSGIGTGQOPAKKRLFGGTGDSSEYVDPQPLGEPATPAA 197
DB      1    TAPGKRPVEQSPOEDSSSGIGTGQOPAKKRLFGGTGDSSEYVDPQPLGEPATPAA 60
QY      198  VGPTTMA SGGA PMADNNEGADGVGNASGNWCHDSTWLGDRVITTSRTTALPTYNNHLY 257
DB      61  VGPTTMA SGGA PMADNNEGADGVGNASGNWCHDSTWLGDRVITTSRTTALPTYNNHLY 120
QY      258  KQISSASTGASNDNHFGYSTPMGYFDENRPHCHFSPPDMQRLINNNMGFRPKLNFELF 317
DB      121  KQISSASTGASNDNHFGYSTPMGYFDENRPHCHFSPPDMQRLINNNMGFRPKLNFELF 180
QY      318  NIQVEKTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQCLPPPADVFMIPQY 377
DB      181  NIQVEKTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQCLPPPADVFMIPQY 240
QY      378  YLTANNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYFEEVPHSSVAHSQSLDRLMNP 437
DB      241  YLTANNGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYFEEVPHSSVAHSQSLDRLMNP 300
QY      438  LIDQYLYLNRTONQSGSAQNKDILFSRGS PAGMSVOPKMWLPFCYRQQRVSXTKTDNN 497
DB      301  LIDQYLYLNRTONQSGSAQNKDILFSRGS PAGMSVOPKMWLPFCYRQQRVSXTKTDNN 360
QY      498  NSNTWTGASKTNLNGRESIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDN 557
DB      361  NSNTWTGASKTNLNGRESIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDN 420
QY      558  VMIDDEEIKATNVATEREGTVA VNFQSSSTDPA TGDVHAMGALPGVMQDRVYLOGP 617
DB      421  VMIDDEEIKATNVATEREGTVA VNFQSSSTDPA TGDVHAMGALPGVMQDRVYLOGP 480
QY      618  IWAKI PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPAPAEFSATKFASTIYOYSTG 677
DB      481  IWAKI PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPAPAEFSATKFASTIYOYSTG 540
QY      678  QVSVEIEMELQKENS KRNMPVQYTSNVAKSANDFTVNNGLYTEPRPIGTRYLTRPL 736
DB      541  QVSVEIEMELQKENS KRNMPVQYTSNVAKSANDFTVNNGLYTEPRPIGTRYLTRPL 599
```

```

RESULT 8
US-09-807-802A-17
; Sequence 17, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; FILE REFERENCE: GMPN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
```

```

; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 534
; TYPE: PR
; ORGANISM: AAV-1
US-09-807-802A-17
```

```

Query Match      72.9%; Score 2906; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 2e-246;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      203  MASGGGAPMADNNGADGVGNASGNWCHDSTWLGDRVITTSRTTALPTYNNHLYKQISS 262
DB      1    MASGGGAPMADNNGADGVGNASGNWCHDSTWLGDRVITTSRTTALPTYNNHLYKQISS 60
QY      263  ASTGASNDNHFGYSTPMGYFDENRPHCHFSPPDMQRLINNNMGFRPKLNFELFNIQY 322
DB      61  ASTGASNDNHFGYSTPMGYFDENRPHCHFSPPDMQRLINNNMGFRPKLNFELFNIQY 120
QY      323  EVTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQCLPPPADVFMIPQYGLTILN 382
DB      121  EVTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQCLPPPADVFMIPQYGLTILN 180
QY      383  NGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYFEEVPHSSVAHSQSLDRLMNP LIDQY 442
DB      181  NGSOAVGRSSFYCLEYFSPQMLRTGNNFTFSYFEEVPHSSVAHSQSLDRLMNP LIDQY 240
QY      443  LYINNRTONQSGSAQNKDILFSRGS PAGMSVOPKMWLPFCYRQQRVSXTKTDNNSNFT 502
DB      241  LYINNRTONQSGSAQNKDILFSRGS PAGMSVOPKMWLPFCYRQQRVSXTKTDNNSNFT 300
QY      503  WTGASKTNLNGRESIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDNVMTD 562
DB      301  WTGASKTNLNGRESIINPGTAMASHKODEDKFFPMGCVMI FGKESAGASNTALDNVMTD 360
QY      563  EEEIKATNPVATEREGTVA VNFQSSSTDPA TGDVHAMGALPGVMQDRVYLOGP IWAKI 622
DB      361  EEEIKATNPVATEREGTVA VNFQSSSTDPA TGDVHAMGALPGVMQDRVYLOGP IWAKI 420
QY      623  PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPAPAEFSATKFASTIYOYSTGOVSE 682
DB      421  PHTDGHFHPSPLMGFGGLKNPPQILIKNTVPANPAPAEFSATKFASTIYOYSTGOVSE 480
QY      683  IEMELOKENS KRNMPVQYTSNVAKSANDFTVNNGLYTEPRPIGTRYLTRPL 736
DB      481  IEMELOKENS KRNMPVQYTSNVAKSANDFTVNNGLYTEPRPIGTRYLTRPL 534
```

```

RESULT 9
US-10-038-972A-14
; Sequence 14, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/3696US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 538
; TYPE: PR
; ORGANISM: adeno-associated virus 2 VP2 capsid protien
US-10-038-972A-14
```

```

Query Match      69.2%; Score 2759.5; DB 2; Length 598;
Best Local Similarity 82.8%; Pred. No. 1.8e-233;
```

```
Matches 495; Conservative 43; Mismatches 59; Indels 1; Gaps 1;
QY 139 APEKRRPVEOSPOEDSSSGIGKTGOQPAKRLNFGQTGDSVDPQPIGEPPATPAAY 198
Db 2 APEKRRPVEHSPEVDPSSSGTGKAGQQPAKRLNFGQTGADSVDPQPIGEPPAASGL 61
QY 199 GPTTASGGAPADNNEGADGVGNASGNHCHSTWLGDRVITTSRTTALPTYYNNHLYK 258
Db 62 GTTMTATGSGAPADNNEGADGVGNASGNHCHSTWLGDRVITTSRTTALPTYYNNHLYK 121
QY 259 QISSASGASNDNHFGYSTPMGCFPDNRFCHFSPPDMORLIINNMGFFPKLNFELFN 318
Db 122 QISSQS-GASNDNHFGYSTPMGCFPDNRFCHFSPPDMORLIINNMGFFPKLNFELFN 180
QY 319 IQYKEVTTNDGVTTIANNLTSTVQVSDSEYQLPYVLGSAHQCLPEPPADVFMIPOGY 378
Db 181 IQYKEVTTNDGVTTIANNLTSTVQVSDSEYQLPYVLGSAHQCLPEPPADVFMIPOGY 240
QY 379 LTIANGSOAVGRSSFFCLEFPPSQMLRTGNFTFSYFEFVPHSSVAHSQSLDRLMPL 438
Db 241 LTIANGSOAVGRSSFFCLEFPPSQMLRTGNFTFSYFEFVPHSSVAHSQSLDRLMPL 300
QY 439 IDOYLXNKTQNSGSAQKDLFSRGSFAGMSVOPQKMLPGCYRQORVSKTIDNN 498
Db 301 IDOYLXNKTQNSGSAQKDLFSRGSFAGMSVOPQKMLPGCYRQORVSKTIDNN 360
QY 499 SNETGASRYNLNGRESIINPGTAMASHKDEDEKFFPMSGWI FGESAGASNTALDNV 558
Db 361 SEVSMGTATKYHLNGRSLVNPGRAMASHKDEDEKFFPMSGWI FGESAGASNTALDNV 420
QY 559 MIDEDEIKATNPVATRFCTVAVNPFSSSTDPATGVAHMGALPGVWMDRDVYLOGPI 618
Db 421 MIDEDEIKATNPVATRFCTVAVNPFSSSTDPATGVAHMGALPGVWMDRDVYLOGPI 480
QY 619 WAKIPHTDGHFSPPLMGFGGLKNPPOILIKTTPVPANPAESATKPFSTOYSTGQ 678
Db 481 WAKIPHTDGHFSPPLMGFGGLKNPPOILIKTTPVPANPAESATKPFSTOYSTGQ 540
QY 679 VSVEIEMELQKENSKRNPPEVOYTSNYKASANDFTVNNGLYTEPRPIGTRYLTRPL 736
Db 541 VSVEIEMELQKENSKRNPPEVOYTSNYKASANDFTVNNGLYTEPRPIGTRYLTRPL 598

RESULT 10
US-09-532-594B-4
; Sequence 4, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorint, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532.594B
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP1
US-09-532-594B-4

Query Match 62.3%; Score 2486.5; DB 2; Length 734;
Best Local Similarity 63.5%; Pred. No. 2.5e-209;
Matches 475; Conservative 81; Mismatches 161; Indels 31; Gaps 10;
```

```
QY 4 DGYLPDWLENDLSEGIREWMDLKRGAPKPKRANOXKDDDRGLVLPGYKYLGPENGLDKE 63
Db 3 DGYLPDWLENDLSEGIREWMDLKRGAPKPKRANOXKDDDRGLVLPGYKYLGPENGLDKE 62
QY 64 PVNAADAAALHEDKAYDQOLKAGDNPYLRYNHADAFOERLQEDTSFGCNLAGAVFOAKK 123
Db 63 PVNAADAAALHEDKAYDQOLKAGDNPYLRYNHADAFOERLQEDTSFGCNLAGAVFOAKK 122
QY 124 RVLEPGLVEBAGKATPKKKRPVEGSPPOEDSSSGIGKTGOQPAKRLNFGQTGDSV 182
Db 123 RVLEPGLVEBAGKATPKKKRPVEGSPPOEDSSSGIGKTGOQPAKRLNFGQTGDSV 182
QY 183 PDQPIGEPPATPAAYGPTTASGGAPADNNEGADGVGNASGNHCHSTWLGDRVITTT 242
Db 183 PDQPIGEPPATPAAYGPTTASGGAPADNNEGADGVGNASGNHCHSTWLGDRVITTT 236
QY 243 STRTWLPTYYNNHLYKRLGE-SSASGASNDNHFGYSTPMGCFPDNRFCHFSPPDMORLI 302
Db 237 STRTWLPTYYNNHLYKRLGE-SSASGASNDNHFGYSTPMGCFPDNRFCHFSPPDMORLI 292
QY 303 NNMGFFPKLNFELFIQYKEVTTNDGVTTIANNLTSTVQVSDSEYQLPYVLGSAHQCL 362
Db 293 NNMGFFPKLNFELFIQYKEVTTNDGVTTIANNLTSTVQVSDSEYQLPYVLGSAHQCL 352
QY 363 LPEPPADVFMIPOGY---LTIANGSOAVGRSSFFCLEFPPSQMLRTGNFTFSYFEFV 419
Db 353 LPEPPADVFMIPOGY---LTIANGSOAVGRSSFFCLEFPPSQMLRTGNFTFSYFEFV 412
QY 420 PFHSSVAHSQSLDRLMPLIDOYLXNKTQNSGSAQKDLFSRGSFAGMSVOP 475
Db 413 PFHSSVAHSQSLDRLMPLIDOYLXNKTQNSGSAQKDLFSRGSFAGMSVOP 469
QY 476 KMLPBCRTRQORVSKTIDNNNSFTWAGS---KY---NLNGRESIINPGTAMASHK 528
Db 470 KMLPBCRTRQORVSKTIDNNNSFTWAGS---KY---NLNGRESIINPGTAMASHK 527
QY 529 DDEDEKFFPMSGWI FGESAGASNTALDNVIMIDEDEIKATNPVATRFCTVAVNPFSSS 588
Db 528 DDEDEKFFPMSGWI FGESAGASNTALDNVIMIDEDEIKATNPVATRFCTVAVNPFSSS 586
QY 589 TDPATGVAHMGALPGVWMDRDVYLOGPIWAKIPHTDGHFSPPLMGFGGLKNPPOIL 648
Db 587 TDPATGVAHMGALPGVWMDRDVYLOGPIWAKIPHTDGHFSPPLMGFGGLKNPPOIL 646
QY 649 IKTTPVPANPAESATKPFSTOYSTGQVSVEIEMELQKENSKRNPPEVOYTSNYAKS 708
Db 647 IKTTPVPANPAESATKPFSTOYSTGQVSVEIEMELQKENSKRNPPEVOYTSNYAKS 706
QY 709 ANVDFTVNNGLYTEPRPIGTRYLTRPL 736
Db 707 ANVDFTVNNGLYTEPRPIGTRYLTRPL 734

RESULT 11
US-10-038-972A-15
; Sequence 15, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36960US
; CURRENT APPLICATION NUMBER: US/10/038.972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 533
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VP3 capsid protien
US-10-038-972A-15
```

```
Query Match      62.2%; Score 2481.5; DB 2; Length 533;
Best Local Similarity 83.3%; Pred. No. 4.1e-209;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;

Qy 203 MASGGA PMADNNNGADGVNAGNMHCDSTWMLGDRVTTSTRTMALPTYNHLYKOISS 262
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MATSSGAPMADNNNGADGVNAGNMHCDSTWMLGDRVTTSTRTMALPTYNHLYKOISS 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 263 ASTGASNDNHFGYSTPWGYPEDFNRFCHFSPRDMQRLINNNGWFRPRLNFKLNIQVK 322
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 QS -GASNDNHFGYSTPWGYPEDFNRFCHFSPRDMQRLINNNGWFRPRLNFKLNIQVK 119
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 323 EVTTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQCLPPRPADVPMIPOYGYLTIN 382
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 EVTQNDGVTTIANNLTSTVQVFTDSEYQLPYVLGSAHQCLPPRPADVPMIPOYGYLTIN 179
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 383 NGSOAVGRSSFYCLEYFPSSOMLRTGNNTFSYTFEEVEFHSSVYAHOSLIDMLNPLIDQY 442
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 NGSOAVGRSSFYCLEYFPSSOMLRTGNNTFSYTFEDVFHSSVYAHOSLIDMLNPLIDQY 239
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 443 LYYLNRTONQSGSAQNKDILLFSRGSFAGMSVOPKMWLPGPCYRQORVSKTIDNNNSNFT 502
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 LYYLSRNTPTSGTTOGRLOFSQAGASDIRQSRNMLPGCYRQORVSKTIDNNNSNFT 299
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 503 WTGASKTNLNGRESIINPTAMASHKDEDEKFFPMGSMVITGKESAGASNTALDNMTD 562
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 WTGATKTHLNGRDLVNPGRMASHKDEEKFPOSQVLIKGKQSEKTNVDIEKVMITD 359
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 563 EEEIKANPVATEFRGVAVNVFOSSSTDPAQGVHAMQALRCGMWQORBDVYLOQPIYAKI 622
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 EEEIRTNPVATEBOYGSVSTNLQGRNROAATADVNTQCVLPQMWQODVDVYLOQPIYAKI 419
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 623 PHTDGHFHPSPBLMGFGGLKNPPPOILLKNTVPANPAPAFSATKFASTTOYSTGOVSVE 682
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 420 PHTDGHFHPSPBLMGFGGLKNPPPOILLKNTVPANPAPAFSATKFASTTOYSTGOVSVE 479
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 683 IEMWLOKENSKRNPVEQVYTSNYAKSANDVFTVNNGLYTEPRPIGTRYLTRPL 736
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 480 IEMWLOKENSKRNPVEIQYTSNYAKSVVDFVTDVNGVSEPRDIGTRYLTRPL 533
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
US-09-533-427-4
; Sequence 4, Application US/09533427
; Patent No. 6855314
; GENERAL INFORMATION:
; APPLICANT: Chlorini, John
; APPLICANT: Kotlin, Robert M.
; APPLICANT: Safier, Brian
; APPLICANT: Davidson, Elizabeth
; APPLICANT: Zahner, Joseph
; TITLE OF INVENTION: AAVS VECTOR FOR TRANSDUCING BRAIN CELLS AND LUNG CELLS
; FILE REFERENCE: 14014.032302
; CURRENT APPLICATION NUMBER: US/09/533.427
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/No. 6855314e =
; OTHER INFORMATION: synthetic construct
US-09-533-427-4

Query Match      55.7%; Score 2220; DB 2; Length 724;
Best Local Similarity 58.6%; Pred. No. 6.4e-166;
Matches 431; Conservative 81; Mismatches 198; Indels 26; Gaps 9;

Qy 8 PDMLENDLSEGIRBWMPLKPGAPKPKXANQKQDDGRGIVGVYGLGFENLGDGEPPVNA 67
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 8 PDMLEF-VGBGLRFFLGLGAGPPKPKPNQHQDARGLVLGIVNYLGGNGLDGGEPPVR 66
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Qy 68 ADAALAEHDAKYDOQLKAGDNPYLRYNHADAEPQERLOEDTSPFCGNGIGRAVFOAKRYLE 127
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 67 ADEVARAEHDISSYNQLEAGDNPYLYKYNHADAEPQERLADTSPFCGNGIGRAVFOAKRYLE 126
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 128 PLGIVGEGAKTAPKPKRPEVQSPQEPDSSSGIGKTGOOPAKKRLNFGTGTGSESPDPQ 187
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 127 PFGIVGEGAKTAPKPKRIDHFPKPKART-----EEDSKP---STSSDAEAGPSSQ 176
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 188 LGSEPPATPA-AVGPTTMAAGGAPMADNNNGADGVNAGNMHCDSTWMLGDRVTTSTRT 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 177 QLOIPIAPQASSLGADTMSAGGGGFLGDNNGADGVNAGNMHCDSTWMLGDRVTTSTRT 236
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 247 WALPTYNHLYKOISSASTGASNDNHFGYSTPWGYPEDFNRFCHFSPRDMQRLINNNG 306
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 237 WVLPSYNNHQYREIKSGSVDSNANAYGYSTPWGYPEDFNRFHSHWSPRDMQRLINNNG 296
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 307 FRPKRLNFKLNIQKEVYTTNDGVTTIANNLTSTVQVFSDSYQLPYVLGSAHQCLPPF 366
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 297 FRPKRLNFKLNIQKEVYTTNDGVTTIANNLTSTVQVFTDDYQLPYVVGNGTEGCLPAF 356
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 367 PADVFMIPOYGYLTIN--NGSOAVGRSSFYCLEYFPSSOMLRTGNNTFSYTFEEVEFHSS 424
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 357 PPQVFTLPQYGYATLNDNTNPTFRKSSFCLEIFPSSMLRTGNNTFEETVFEVYFHHSS 416
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 425 YAHOSLIDRLNPLIDQYLYLNRTONQSGSAQNKDILLFSRGSFAGMSVOP-KMWLPGPC 483
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 417 FAPSQNLFKLANPLVDQYLVRFVSTNNTGQVQFNKUL-----AGRAYNTYKQWFPQPM 469
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 484 YRQORVSKTIDNNNSNFTWTGASKTNLNGRESIINPTAMASHKDEDEKFFPMGSMVIF 543
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 470 GRTQGMWLGSGVNRASVASAFATNTRMELEGASVOPPPQNGMTNNLQGSNTYALENTAIF 529
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 544 GKESAGASNTAL---DNVMTDEBEIKATNPVATERFCTVANFOSSSTDPAQGVHAMG 600
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 530 NSQAPNPGTTATYLEGNNLITSESETQPVNRVAIVNGQAMKTNQOSTTAPATGYINQOE 589
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 601 ALPGWQVQDRDVYLOQPIYAKIPHTDGHFHPSPBLMGFGGLKNPPPOILLKNTVPANPAP 660
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 590 IVPGSVMNERDVYLOQPIYAKIPETGAHFHSPBLMGFGGLKPPPMILKNTVPAGN-IT 648
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 661 EFSATKFASTTOYSTGOVSVEIEMWLOKENSKRNPVEQVYTSNYAKSANDVFTVNNGL 720
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 649 SFSQVPPVSSFTTOYSTGOVIVEMEMELKENSKRNPVEIQYTSNNYNDPQVDFAPADSTGE 708
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 721 YTEPRPIGTRYLTRPL 736
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 709 YTRTRPIGTRYLTRPL 724
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
US-09-717-789C-4
; Sequence 4, Application US/09717789C
; Patent No. 6984517
; GENERAL INFORMATION:
; APPLICANT: Chlorini, John
; APPLICANT: Kotlin, Robert M.
; APPLICANT: Safier, Brian
; TITLE OF INVENTION: AAVS VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.032303
; CURRENT APPLICATION NUMBER: US/09/717.789C
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: PCT/US99/11958
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/087, 029
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```



```

1  TITLE OF INVENTION: AAVS VECTOR FOR TRANSDUCING BRAIN CELLS AND LUNG CELLS
2
3  FILE REFERENCE: 14014,023302
4
5  CURRENT APPLICATION NUMBER: US/09/533,427
6
7  CURRENT FILING DATE: 2000-03-22
8
9  NUMBER OF SEQ ID NOS: 23
10
11  SOFTWARE: FastSeq for Windows Version 3.0
12
13  SEQ ID NO 5
14
15  LENGTH: 588
16
17  TYPE: PRT
18
19  ORGANISM: Artificial Sequence
20
21  FEATURE:
22
23  OTHER INFORMATION: Description of Artificial Sequence:/No. 6855314e =
24
25  OTHER INFORMATION: Synthetic construct
26
27  US-09-533-427-5

```

Query Match	42.6%	Score 1700.5	DB 2	Length 588
Best Local Similarity	55.0%	Pred No. 2.1e-140		
Matches 332	Conservative 64	Mismatches 173	Indels 35	Gaps 7

Qy	140	PGKKRPPEQSPQEPDSSSGIGKGTQGP	PAKKRLINFGQTGSESEVPQEPGLSEPPAPA-AV	198
Dd	13	PKKKARTBEDSKRPSTSS-----	DAEAGPSSQOQIPIAOPASSL	52
Qy	199	GPTTMAAGGAPMADNNEGADGVNAGNMWCDSTWLGADRVITTSRTKTALEPTYNHLK		256
Dd	53	GADTMASAGGGGPIGADNNQGADGVNAGNMWCDSTWMEGRVYTKSTRITWVLPYSNNHQYR		112
Qy	259	QISASATGASNDNHHYGYSTPMGYEPDPAHFEHCFSPROMQRLINNWGFRPRKRLNFKLN		312
Dd	113	EISGSGVDSGNAAYGYSTPMWGYEPDNFPHSHMSRDMQRLINNYWGFRRPSRLKVLKN		178
Qy	319	IQVKEVTTNDGVTITANNLTSTVQVPSDSEYQLPYLGASAGGCLPPPADVFMIPQYGY		376
Dd	173	IQVKEVTVQDSTTIANNLTSTVQVFTDDYQLPYVNGTGTGCLPAPRPVFTLPQYGY		232
Qy	379	LITN--NGSOAVGRSSFYCLETFPSSOMLRGTNNFTFSYTFEEVPHSSVYAHOSQIDRLN		436
Dd	233	ATINRDNTEPTEPSSFFCLIEFYPPSMLTGTNNFETVYFEEVPHSSFAFSOMLFTKAN		292
Qy	437	PLIDQYLXYLNRQONGSGAKONDLIFSRGSPAGMSVQR-KWMLGPCYRQORVSKTKTD		492
Dd	293	PLVDQIHYRVSITNNNGVQFENKUL-----AGRYANTYKMFEPGPRGRTQGMKLGSQV		348
Qy	496	NNNSNFTWTGASKYNLNGRESIINPCTAMASHKDEDEKFPMSGVMIFGKESAGASNTAL		552
Dd	346	NRASVASFATITNMELEGASYQPPQPNQGMATNNLQGSNTVALENTMFINSQPANDGTAT		402
Qy	556	---DNMIIIDDEEIKATNPATATPATERPCTVAVNFQSSSTDPATGDDVAMGALQGMWQDRDV		612
Dd	406	YLEGNMLITSESETQPVNRYAVVNGQAMATNNQSSSTPATCTVYMLQEIVGSGVMEMBD		462
Qy	613	YLOGPIWAKIPHTDGHFHPSPMLGGGLGNPPOLIKNTPVANPAPAEFSATVPASFIT		672
Dd	466	YLOGPIWAKIPETGALFHPSPAMGGGLGHPHPMLIKNTPVGN-ITISFDVPSSSIT		522
Qy	673	QYSTGVQSVIEIEMLOKENSKRANPEVOYTSYVAKSANDFTVDNNGLYTEBPRIGRYL		732
Dd	525	QYSTGVQVTEMEMELCKENSKRANPEIQYTNNYNDPQFVDFAFDSTGEYRTTRPIGRYL		582
Qy	733	TRPL 736		
Dd	585	TRPL 588		

Search completed: June 12, 2006, 12:54:12
Job time : 55 secs

THIS PAGE BLANK (USPTO)